

Optimal search for a moving target with the option to wait

Citation for published version (APA):

Flesch, J., Karagozoglu, E., & Perea y Monsuwé, A. (2007). *Optimal search for a moving target with the option to wait*. METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 051 <https://doi.org/10.26481/umamet.2007051>

Document status and date:

Published: 01/01/2007

DOI:

[10.26481/umamet.2007051](https://doi.org/10.26481/umamet.2007051)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

János Flesch, Emin Karagozoglu,
Andrés Perea

Optimal Search for a Moving Target with the
Option to Wait

RM/07/051

JEL code: C44, C61, C72



Maastricht research school of Economics
of **TE**chnology and **OR**ganizations

Universiteit Maastricht
Faculty of Economics and Business Administration
P.O. Box 616
NL - 6200 MD Maastricht

phone : ++31 43 388 3830
fax : ++31 43 388 4873

Optimal Search for a Moving Target with the Option to Wait

János Flesch*, Emin Karagozoglu† and Andrés Perea‡

November 29, 2007

Abstract

We investigate the problem in which an agent has to find an object that moves between two locations according to a discrete Markov process (see Pollock, 1970). At every period, the agent has three options: searching left, searching right, and waiting. We assume that waiting is costless whereas searching is costly. Waiting can be useful because it could induce a more favorable probability distribution over the two locations next period. We find an essentially unique (nearly) optimal strategy, and prove that it is characterized by two thresholds (as conjectured by Weber, 1986). We show, moreover, that it can never be optimal to search the location with the lower probability of containing the object. The latter result is far from obvious and is in clear contrast with the example in Ross (1983) for the model without waiting.

We also analyze the case of multiple agents. This makes the problem a more strategic one, since now the agents not only compete against time but also against each other in finding the object. We find different kinds of subgame perfect equilibria, possibly containing strategies that are not optimal in the one-agent case. We compare the various equilibria in terms of cost-effectiveness.

JEL Classification Codes: C44, C61, C72

Keywords: Search for a moving target, waiting, Markovian dynamics, subgame perfect equilibrium.

*Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. E-mail: J.Flesch@KE.unimaas.nl

†Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. E-mail: E.Karagozoglu@ALGEC.unimaas.nl

‡Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. E-mail: A.Perea@KE.unimaas.nl

1 Introduction

1.1 Description of the Problem without Waiting

In 1970, Pollock considered the following search problem: there is an object moving across two locations labeled as L (left) and R (right) according to a discrete time Markov process (see Figure 1).

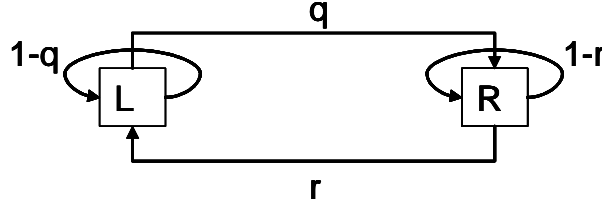


Figure 1: Search For a Moving Target

Hence, if the object currently is at L , it will move to R with probability q . If the object currently is at R , it will move to L with probability r . The probabilities q and r are called the *transition probabilities*. The probability that the object is initially at L is \bar{p} , and hence $1 - \bar{p}$ is the probability that the object initially is at R .

The periods at which the agent can search for the object are denoted by $t = 1, 2, 3, \dots$. As soon as the agent finds the object, the process stops. If the object has not yet been found, the agent has two possible actions at any given period: searching left (L) and searching right (R), which have costs c_L and c_R , respectively. Assume that the object is overlooked with probability α_L if it is located at L and the agent searches L , and overlooked with probability α_R if it is located at R and the agent searches R . The objective is to minimize the total expected searching cost needed for finding the object.

1.2 Overview of the Literature

Ross (1983) conjectured that there exists an optimal searching strategy of the following form: search L at period t if and only if $p_t \geq \pi$, where p_t is the probability that the object is located at L at period t . Here, π is a threshold, which depends on the parameters of the model (i.e., q, r , and \bar{p}). Although this conjecture was intuitive, it has never been proven in full generality until now.

Pollock (1970) showed that the conjecture is valid for the special case without overlooking (i.e., $\alpha_L = \alpha_R = 0$) in the discrete time model. He also computed the threshold $\tilde{\pi}$ for every possible configuration of parameters. Weber (1986) proved Ross' conjecture for a continuous time Markov process, including the possibility of overlooking. He also conjectured that for the case with waiting, the optimal strategy can be characterized by two thresholds. Later, MacPhee and Jordan (1995) proved Ross' conjecture for a discrete time Markov process with the possibility of overlooking, but only for some parameter configurations.

Assaf and Sharlin-Bilitzky (1994) investigated a continuous time version of the model in which the agent, in addition, can specify the searching effort. The main result of the paper is to describe the optimal searching strategy in this model. Kan (1977) investigated the search problem above with the additional option to stop searching. The paper characterized optimal strategies for specific classes of transition probabilities.

1.3 Our Contribution

Our paper extends the original search problem by introducing the option to wait.¹ That is, at every period the agent has now three options: searching left (L), searching right (R), and waiting (0). We assume that waiting is costless whereas searching is costly. This extension is motivated by the fact that an agent can benefit from waiting by getting a more favorable probability distribution over the two locations next period.

We assume that the costs of searching left and right are equal (normalized to 1) and the overlooking probabilities are 0. If the agent finds the object he receives a prize P . Provided the prize is high enough, this gives the agent an incentive to find the object with probability 1. Without this prize, it would be optimal to simply wait forever.

We prove that, if $P \geq 2$ and the sum of the transition probabilities (i.e., $q + r$) is at least 1, or $q = r$, then there exists an optimal strategy of the following form: search R if $p_t \in [0, \pi_1]$, wait if $p_t \in (\pi_1, \pi_2)$ and search L if $p_t \in [\pi_2, 1]$. (Recall that p_t denotes the probability that the object is at location L at period t , given it has not been found yet.) That is, this optimal strategy is given by two thresholds, π_1 and π_2 , as conjectured by Weber (1986). Moreover, this optimal strategy is unique up to the choices at π_1 and π_2 .

On the other hand, if $P \geq 2$ and $0 < q + r < 1$ and $q \neq r$, we prove that for every $\varepsilon > 0$ there exists an ε -optimal strategy of the form described above. By an ε -optimal strategy we mean that the agent cannot improve his expected utility by more than ε by switching to another strategy. In this case, an optimal strategy need not exist since there are situations in which waiting for one more period always provides more favorable probabilities p_t and $1 - p_t$ for the next period. In these situations, the agent wants to postpone searching as long as possible.

We show that our thresholds π_1 and π_2 have an interesting relationship with the threshold $\tilde{\pi}$ as derived by Pollock for the model without waiting: It turns out that $\tilde{\pi}$ is always in between π_1 and π_2 . Consequently, if the agent searches a location in the model with waiting, he would search the same location in the model without waiting. Although this result is not unexpected, it is not immediately clear why it should be the case.

In the model without the option to wait, Ross (1983) has shown that it is possible that searching the location with the lower probability of containing the object is

¹Note that waiting is different from stopping analyzed in Kan (1977). The agent can continue searching after waiting, whereas stopping means no more search can be done.

optimal. The reason is that searching this location may induce a very favorable probability distribution for the location of the object next time, in case it is not found. We show that this is no longer possible in our model: the agent either searches the location with the higher probability of containing the object, or he waits. So, even if the lower probability is close to $\frac{1}{2}$, and searching the location with the lower probability would provide the agent with complete certainty next period, he would still not search that location. This fact is far from obvious.

As another extension, we analyze the case of multiple agents. This extension makes the problem a more strategic one, since now the agents not only compete against time but also against each other in finding the object. In that sense, this extension embeds the two frequently studied objectives in the literature: minimizing the expected cost of finding the object and maximizing the probability of finding the object in a given time.² In particular, we assume that if agent i is the first one to find the object, possibly at the same time as some of his opponents, then his utility would be equal to some prize $P \geq 2$ (for finding the object first) minus his total searching cost. If he does not find the object first, he will not receive the prize P but still incurs the cost of searching. Hence, each agent wants to minimize his expected search cost and also maximize his probability of finding it in a given time.

We show that, if $P \geq 2$, there are multiple subgame perfect equilibria in this search game. There is a symmetric subgame perfect equilibrium in which every agent follows the single-agent optimal strategy. There is another symmetric subgame perfect equilibrium, however, in which every agent searches the location with the higher probability of containing the object. If the prize is high enough, there is even a symmetric subgame perfect equilibrium in which every agent always searches the location with the lowest probability of containing the object. We show that the equilibrium in which every agent plays his one-person optimal strategy Pareto dominates all other symmetric equilibria.

1.4 Road Map

Section 2 of this paper introduces the necessary notation and presents our main theorem along with the values of the thresholds π_1 and π_2 . Section 3 introduces the value function and analyzes its properties, which we use in the proof of the main theorem. Section 4 discusses all different types of Markov processes. Section 5 contains preparatory lemmas for the proof of the main theorem. Section 6 contains two lemmas which prove the existence of the thresholds for all possible types of Markov processes. In Section 7, we compute the values of the thresholds π_1 and π_2 (see Table 1). In Section 8, we compare our optimal strategy with the optimal strategy for the model without waiting, as derived by Pollock (1970). In Section 9 we introduce the model with multiple agents and analyze the equilibria of the search game. Section 10 provides a discussion of the assumptions of the model, as well as a comparison of the results from different models.

²See, for instance, Kan (1977) or Pursiheimo (1978).

2 Main Theorem

2.1 Optimal and ε -Optimal Strategies

A *strategy* γ is defined as a sequence $(\gamma_t)_{t \in \mathbb{N}}$, where $\gamma_t \in \{L, R, 0\}$ describes the action to be taken at period t if the object has not been found yet.³ The expected utility induced by a strategy γ is denoted by $U(\gamma)$, while the expected searching cost is called $C(\gamma)$. Thus, $U(\gamma) = \mu(\gamma)P - C(\gamma)$, where $\mu(\gamma)$ is the probability of finding the object.

We now mention some specific classes of strategies. We call a strategy γ *successful* if $\mu(\gamma) = 1$. In that case, $U(\gamma) = P - C(\gamma)$. For every period t , let $\mu_t(\gamma)$ be the probability of finding the object at period t , given that it has not been found yet. Also, let $C_t(\gamma)$ be the cost at period t (so, $C_t(\gamma) = 1$ if the agent searches, and 0 if he waits). We call the strategy γ *loss-free* if $\mu_t(\gamma)P - C_t(\gamma) \geq 0$ for every period t , meaning that at every period the expected utility is non-negative. Moreover, we call γ *strictly loss-free* if $\mu_t(\gamma)P - C_t(\gamma) > 0$ for every period t at which γ prescribes to search.

Formally, the agent's objective is to find, for a given configuration (\bar{p}, q, r) of initial and transition probabilities, a strategy γ that maximizes the expected utility, $U(\gamma)$. A strategy γ is called *optimal* if $U(\gamma) \geq U(\gamma')$ for all other strategies γ' . For any $\varepsilon > 0$, the strategy γ is called *ε -optimal* if $U(\gamma) \geq U(\gamma') - \varepsilon$ for all other strategies γ' . As we will show below, optimal strategies do not always exist for every configuration of initial and transition probabilities, but for every $\varepsilon > 0$ we can always construct an ε -optimal strategy.

We will show that, if $P \geq 2$, then it is sufficient to consider successful strategies.

Lemma 2.1 *Let $P \geq 2$, and let γ be an arbitrary strategy. Then, there is a successful strategy γ' with $U(\gamma') \geq U(\gamma)$.*

Proof. Suppose that γ is not successful. For every period t , let $z_t(\gamma)$ be the conditional probability that the object will be found from period t on, given that the object has not been found before t . We will show that $z_t(\gamma)$ tends to 0 as $t \rightarrow \infty$.

Note that $z_1(\gamma)$ is simply the probability that the object will eventually be found. So, by assumption, $z_1(\gamma) < 1$. Let $\bar{z}_t(\gamma)$ be the probability that the object will be found before period t . Then,

$$z_1(\gamma) = \bar{z}_t(\gamma) + (1 - \bar{z}_t(\gamma))z_t(\gamma)$$

for all t . Since $\bar{z}_t(\gamma) \rightarrow z_1(\gamma)$ as $t \rightarrow \infty$, and $1 - \bar{z}_t(\gamma) \rightarrow 1 - z_1(\gamma) > 0$ as $t \rightarrow \infty$, we have that $z_t(\gamma) \rightarrow 0$ as $t \rightarrow \infty$.

So, there is some period t with $z_t(\gamma) \leq \frac{1}{P}$. This implies that from period t on the expected continuation utility for γ is at most 0. The reason is that either γ does not search after t , yielding a continuation utility of 0, or searches at least once, yielding a continuation utility of at most $\frac{1}{P}P - 1 = 0$.

³We do not consider mixed or history dependent strategies here, since the agent will not be able to improve upon his cost by turning to such strategies.

Now, let γ' be the strategy that coincides with γ during the first $t - 1$ periods, and which from period t on always searches the location with the highest current probability of containing the object. (If both locations are equally likely, check L). Then, γ' will find the object with probability 1, so is a successful strategy. Suppose that the object has not been found before period t . Then it can be shown by induction on k that the expected searching cost until period $t + k$ will be at most $1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^{k-1}} \leq 2$. Therefore, the expected continuation searching cost for γ' from period t on is at most 2. So, the expected continuation utility for γ' from period t on is at least 0, as $P \geq 2$. But then, $U(\gamma') \geq U(\gamma)$. ■

It thus follows that, if $P \geq 2$, we may restrict ourselves to successful strategies. This is useful since finding an optimal strategy within the class of successful strategies is equivalent to minimizing the expected searching costs.

The prize, P , should be at least 2 to make the agent search. If $P < 2$, the agent may find it optimal to wait forever since the benefit from finding the object does not cover the expected costs of search in some cases. Consider, for instance, the situation where the initial probability \bar{p} is $\frac{1}{2}$, and the transition probabilities q and r are $\frac{1}{2}$ as well. In that case, the expected cost of finding the object would be exactly 2.⁴ Hence, the agent would have no incentive to search at all.

2.2 Our Main Theorem

Our main theorem states that, whatever configuration of initial probabilities and transition probabilities we take, we can always construct a (nearly) optimal strategy of the following kind:

- if the probability that the object is at L is below a threshold π_1 , search R ;
- if the probability that the object is at L is between π_1 and another threshold $\pi_2 \geq \pi_1$, wait;
- if the probability that the object is at L is above π_2 , search L .

The (nearly) optimal strategy is therefore completely characterized by the thresholds π_1 and π_2 . It is possible that the two thresholds coincide. In that case, the agent will never wait.

In the main theorem, let $p_t(\gamma)$ be the probability that at period t the object is at L if strategy γ is being implemented⁵. Note that $p_t(\gamma)$ can easily be computed from γ : If at period $t - 1$ action L has been chosen, then $p_t(\gamma) = r$. Namely, if by choosing L the object has not been found at period $t - 1$, then the object must have been at R at period $t - 1$, and hence will be at L with probability r next period. By

⁴The reason is that at every period, if the object has not yet been found, the probabilities of the object being located at L and R would be exactly $\frac{1}{2}$. So, the expected searching cost will be $\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \dots = 2$.

⁵From now on, whenever we speak about period t , we always assume that the object has not been found before this period.

	π_1	π_2
$q + r = 0$	$\frac{1}{2}$	$\frac{1}{2}$
$0 < q + r < 1$ and $(q + 1)r^2 + q^2r - q \leq 0$ (Area A in Figure 2)	$\frac{q(1+r)}{(q+r)(1+q)}$ (unique p with $V(p, R) = V(\pi^*, L)$)	$\max\{\frac{1}{2}, \frac{r}{q+r} - \frac{\alpha\varepsilon}{2}\}$
$0 < q + r < 1$ and $(q + 1)r^2 + q^2r - q > 0$ (Area B in Figure 2)	$\frac{q}{(q+r)(1+q)-q}$ (unique p with $V(p, R) = V(\pi^*, L)$)	$\max\{\frac{1}{2}, \frac{r}{q+r} - \frac{\alpha\varepsilon}{2}\}$
$q + r = 1$	q	r
$1 < q + r < 2$ and $(1 - q + q^2)r - 2q^2 + q^3 \leq 0$ (Area C in Figure 2)	$\frac{q(1-r)}{(q+r)(1-q)}$ (unique p with $V(p, R) = V(Ap, L)$)	$\frac{r}{q+r}$
$1 < q + r < 2$ and $(1 - q + q^2)r - 2q^2 + q^3 > 0$ (Area D in Figure 2)	$\frac{1-r}{(q+r)(1-q)+1-r}$ (unique p with $V(p, R) = V(Ap, L)$)	$\frac{r}{q+r}$
$q + r = 2$	$\frac{1}{2}$	$\frac{1}{2}$

Table 1: Values for the thresholds π_1 and π_2 when $q \leq r$. If $q < r$ then α is the smallest positive probability amongst \bar{p} , $1 - \bar{p}$, q , $1 - q$, r , and $1 - r$. If $q = r$, then $\alpha = 0$.

a similar reasoning, if at period $t - 1$ action R has been chosen, then $p_t(\gamma) = 1 - q$. If at period $t - 1$ action 0 has been chosen, then

$$p_t(\gamma) = p_{t-1}(\gamma)(1 - q) + (1 - p_{t-1}(\gamma))r.$$

Theorem 2.2 *Let $P \geq 2$. Take a configuration (\bar{p}, q, r) of initial and transition probabilities with $q \leq r$, let $\varepsilon > 0$, and choose the thresholds π_1 and π_2 according to Table 1. Then, the strategy γ given by*

$$\gamma_t = \begin{cases} R, & \text{if } p_t(\gamma) \in [0, \pi_1] \\ 0, & \text{if } p_t(\gamma) \in (\pi_1, \pi_2) \\ L, & \text{if } p_t(\gamma) \in [\pi_2, 1] \end{cases}$$

is an ε -optimal strategy.⁶ Moreover, if $q = r$ or $q + r \geq 1$, then γ is an optimal strategy, unique up to the choices at π_1 and π_2 .

In addition, γ is successful and prescribes to search a location only if the probability that the object is at that location is at least $\frac{1}{2}$. In particular, γ is loss-free.

⁶In fact, the strategy is ε -optimal from any period on.

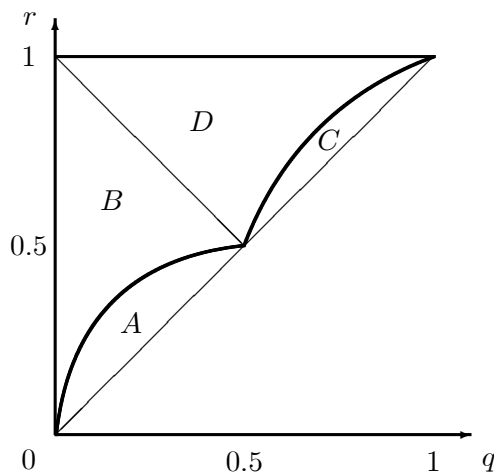


Figure 2: Different areas of (q, r) pairs considered in Theorem 2.2

Figure 2 provides a graphical illustration of the various areas of (q, r) pairs we distinguish in the theorem. Note that the assumption $q \leq r$ is made without loss of generality. If $q < r$ and $q + r \in (0, 1)$, then there are initial probabilities \bar{p} for which no optimal strategy exists. However, we can still guarantee the existence of ε -optimal strategies in this case.

In Table 1, the parts between brackets will only become clear after reading Section 3.2. For our main theorem, any $\tilde{\alpha}$ such that $0 < \tilde{\alpha} < \alpha$ would also do in case $q < r$.

It will follow from our analysis that our thresholds π_1 and π_2 have an interesting relationship with the threshold $\tilde{\pi}$ as computed by Pollock for the model without waiting: It turns out that $\tilde{\pi}$ is always in between π_1 and π_2 . Consequently, if the agent searches a location in the model with waiting, he would search the same location in the model without waiting. Although this result may seem intuitive, it is not immediate.

The proof of Theorem 2.2 is structured as follows. Section 3 introduces the value function and analyzes its properties, which is fundamental for the proof of the main theorem. Section 4 discusses all different types of Markov processes. Section 5 contains preparatory lemmas for the proof of the main theorem. Section 6 contains two lemmas which prove the existence of the thresholds and show that the induced strategies are successful. We also show that these thresholds are unique if $q = r$ or $q + r \geq 1$. Consequently, we prove for these cases the uniqueness of the optimal strategy, up to the choices at the two thresholds. In Section 7, we compute the values of the thresholds π_1 and π_2 (see Table 1), and show that the induced strategy only searches a location if the object is there with probability at least $\frac{1}{2}$.

3 Value Function

3.1 Definition

The key concept in the proof of our main theorem is that of a *value function*. Fix some transition probabilities q and r . Suppose that the object is located at L with probability p . Then, for every strategy γ we denote by $C(p, \gamma)$ the expected cost of finding the object if γ is used. The *value* $V(p)$ is given by

$$V(p) = \inf_{\gamma \text{ successful}} C(p, \gamma)$$

and denotes the minimal expected cost of finding the object if the object is currently located at L with probability p . The function that assigns to every p the number $V(p)$ is called the *value function*. Hence, if a successful strategy γ is optimal for the configuration (\bar{p}, q, r) , then $C(\bar{p}, \gamma) = V(\bar{p})$. Conversely, if $P \geq 2$ and γ is successful with $C(\bar{p}, \gamma) = V(\bar{p})$, then, in view of Lemma 2.1, γ is optimal.

3.2 Basic Properties of the Value Function

For each action $a \in \{L, R, 0\}$, we define the expected cost $V(p, a)$ induced by using action a , given that the agent behaves optimally afterwards. As before, let p denote the probability that the object is currently at L .

When the agent searches L , then with probability p he finds the object and his total cost will be 1, while with probability $(1 - p)$ he cannot find it. In case he cannot find it, he searched already once, and will have a future expected cost of $V(r)$ if he acts optimally from next period on. Therefore, the expected cost of *searching* L can be written as

$$V(p, L) = p \cdot 1 + (1 - p) [1 + V(r)] = 1 + (1 - p)V(r). \quad (1)$$

Similarly, the expected cost of *searching* R is

$$V(p, R) = (1 - p) \cdot 1 + p [1 + V(1 - q)] = 1 + pV(1 - q). \quad (2)$$

Finally, the expected cost of *waiting* is

$$V(p, 0) = V(Ap), \quad (3)$$

where

$$Ap := p(1 - q) + (1 - p)r = (1 - q - r)p + r \quad (4)$$

denotes the probability that the object will be at L in the next period, if the object is currently at L with probability p .

Therefore, the value function satisfies

$$V(p) = \min\{V(p, L), V(p, R), V(p, 0)\}.$$

An action $a \in \{L, R, 0\}$ with $V(p, a) = V(p)$ is called an *optimal action* at p . Moreover, for a given $\varepsilon > 0$ we say that action a is ε -optimal at p if $V(p, a) \leq V(p) + \varepsilon$.

As it can be seen from the equations above, $V(p, L)$, $V(p, R)$ and Ap are all linear in p . Moreover, $V(p, L)$ is strictly monotone decreasing in p , whereas $V(p, R)$ is strictly monotone increasing in p . Combining this with the fact that $V(p, R)$ attains its minimum, which is equal to 1, at $p = 0$ and $V(p, L)$ attains its minimum, which is equal to 1, at $p = 1$, we can conclude that $V(p, L)$ and $V(p, R)$ always intersect at a unique $\bar{\pi} \in (0, 1)$, which is given by the equation

$$\bar{\pi} = \frac{V(r)}{V(1-q) + V(r)}. \quad (5)$$

Hence,

$$\begin{aligned} V(p, L) &> V(p, R) \text{ for all } p \in [0, \bar{\pi}) \\ V(p, L) &= V(p, R) \text{ at } p = \bar{\pi} \\ V(p, L) &< V(p, R) \text{ for all } p \in (\bar{\pi}, 1]. \end{aligned}$$

Figure 3 provides a graphical representation of the functions $V(p, L)$ and $V(p, R)$.

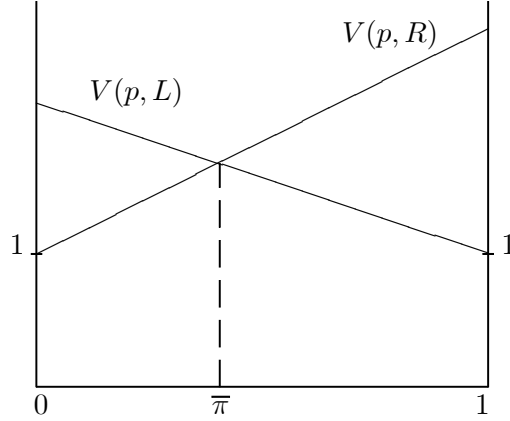


Figure 3: Functions $V(p, L)$ and $V(p, R)$

We denote the invariant distribution(s) of the Markov process by $(\pi^*, 1 - \pi^*)$. In particular, π^* is a solution of the equation

$$A\pi^* = \pi^* \Leftrightarrow (1 - q - r)\pi^* + r = \pi^*.$$

If $q + r > 0$, then this equation has a unique solution which is

$$\pi^* = \frac{r}{q + r}. \quad (6)$$

In the following lemma we list some general properties of the value function that will be useful for the proof of the main theorem.

Lemma 3.1 *The value function $V(p)$ satisfies the following properties;*

- (i) $1 \leq V(p) \leq 2$, and $V(p) = 1 \Leftrightarrow p = 0$ or $p = 1$.
- (ii) $V(p) = \inf_{n \in \{0,1,\dots\}} \min\{V(A^n p, L), V(A^n p, R)\}$.
- (iii) $V(\pi^*) = \min\{V(\pi^*, L), V(\pi^*, R)\}$ if $q + r > 0$.
- (iv) $V(p) \leq V(A^n p) \forall n \in N$, and $V(p) \leq V(\pi^*)$ if $0 < q + r < 2$.
- (v) $V(1 - q) \leq \frac{1}{q}$ if $q > 0$ and $V(r) \leq \frac{1}{r}$ if $r > 0$.

Proof. (i) It is clear that $V(p) \geq 1$ and $V(p) = 1 \Leftrightarrow p = 0$ or $p = 1$. It remains to prove that $V(p) \leq 2$. Consider the strategy γ which, at every period t , searches the location that contains the object with the highest probability (if both locations are equally likely, check L). Then, by the proof of Lemma 2.1, we know that $C(p, \gamma) \leq 2$, yielding $V(p) \leq 2$.

(ii) Since $V(p)$ restricts to successful strategies, the agent can only wait for a finite number of periods in a row, after which L or R has to be chosen.

(iii) This follows from (ii) and the fact that $A^n \pi^* = \pi^*$ for every n .

(iv) By definition, $V(p) \leq V(p, 0) = V(Ap)$. Applying transformation A recursively, we get

$$V(p) \leq V(Ap) \leq V(A^2 p) \leq \dots \leq V(A^{n-1} p) \leq V(A^n p)$$

for every n .

Since $0 < q + r < 2$, the sequence $A^n p$ converges to π^* . By (ii), $V(p)$ is continuous in p , and hence we have that $V(A^n p)$ converges to $V(\pi^*)$. Therefore, $V(p) \leq \lim_{n \rightarrow \infty} V(A^n p) = V(\pi^*)$.

(v) We have $V(1 - q) \leq V(1 - q, R) = 1 + (1 - q)V(1 - q)$. Hence, if $q > 0$, this implies

$$V(1 - q) \leq \frac{1}{q}.$$

Similarly, one can show that $V(r) \leq \frac{1}{r}$, if $r > 0$. ■

4 Five Possible Types of Dynamics

We distinguish 5 different types of dynamics in the Markov process, induced by the transition probabilities q and r .

(i) Absorbing case: $q + r = 0$ i.e. $q = r = 0$.

Object does not move in this case. Function Ap is given by $Ap = p$. Invariant distribution π^* is not unique. In fact, any $\pi^* \in [0, 1]$ is invariant.

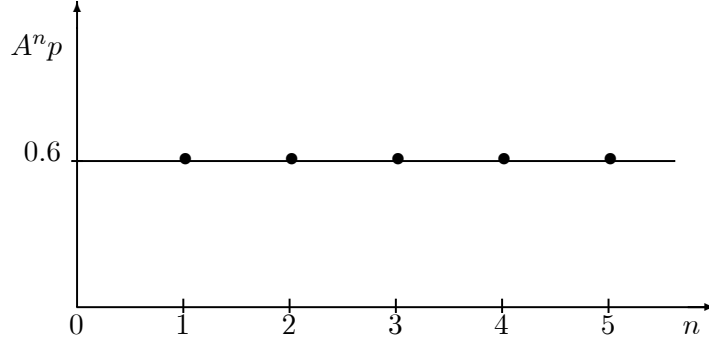


Figure 4: Dynamics for the first five iterations for $p = 0.6$, $q = r = 0$.

(ii) Non-oscillating case: $0 < q + r < 1$.

Function Ap is strictly increasing in p . Invariant distribution π^* is unique. Probabilities Ap and p are always on the same side of π^* . That is, $A^n p$ converges to π^* in a monotonic (and hence non-oscillating) fashion.

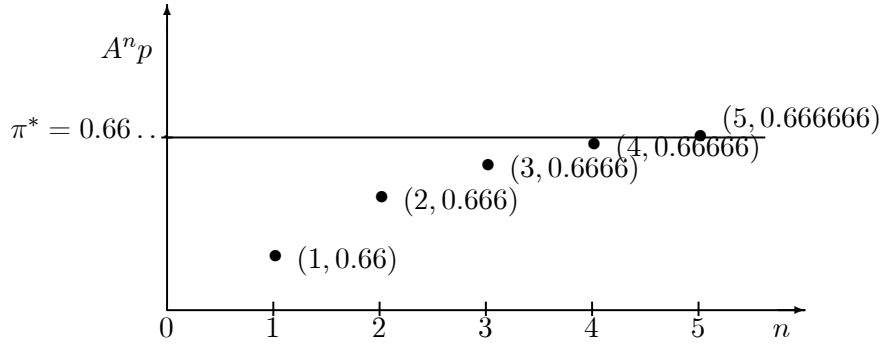


Figure 5: Dynamics for the first five iterations for $p = 0.6$, $q = 0.3$ and $r = 0.6$.

(iii) State independent transitions case: $q + r = 1$.

Function Ap is constant in p . Invariant distribution π^* is unique. The convergence to π^* is immediate. Since $r = 1 - q$, the transitions are independent of the state.

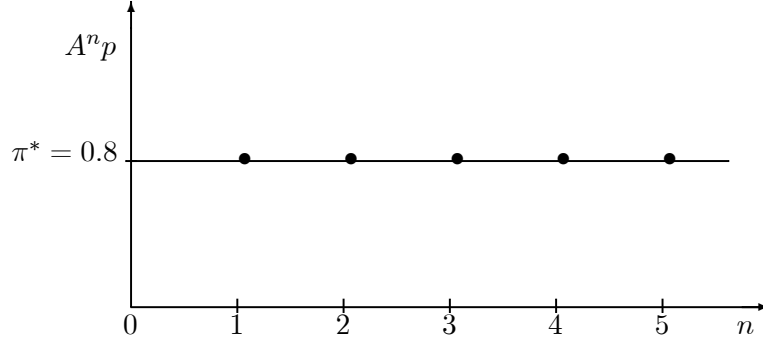


Figure 6: Dynamics for the first five iterations for $p = 0.6$, $q = 0.2$ and $r = 0.8$.

(iv) Oscillating case: $1 < q + r < 2$.

Function Ap is strictly decreasing in p . Invariant distribution π^* is unique. Probabilities Ap and p are always on different sides of π^* . So, $A^n p$ converges to π^* in an oscillating fashion although $|A^n p - \pi^*|$ is a strictly decreasing function of $n \in \mathbb{N}$.

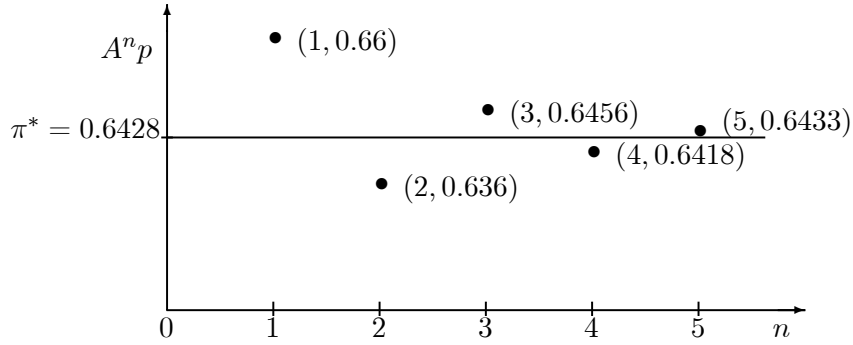


Figure 7: Dynamics for the first five iterations for $p = 0.6$, $q = 0.5$ and $r = 0.9$.

(v) Switching case: $q + r = 2$ i.e. $q = r = 1$.

Function Ap is given by $Ap = 1 - p$. Invariant distribution π^* is unique and equal to $\frac{1}{2}$. By waiting, the probability that the object is at L switches from p to $1 - p$. So, unless $\bar{p} = \pi^*$, there will be no convergence to π^* .

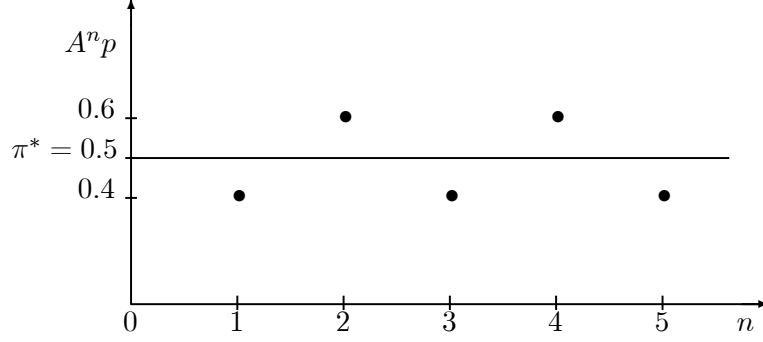


Figure 8: Dynamics for the first five iterations for $p = 0.6$, $q = r = 1$.

5 Preparatory Lemmas

In order to prove our main theorem, we first show that a sequence of optimal actions, under mild conditions, always leads to an optimal strategy.

Lemma 5.1 *Let γ be a strategy which never waits infinitely long. Let $p_t(\gamma)$ be the probability that at period t the object is located at L if the object has not been found yet, and strategy γ is being implemented. Suppose that there is an $\alpha > 0$ such that L is chosen only if $p_t(\gamma) \geq \alpha$, and R is chosen only if $1 - p_t(\gamma) \geq \alpha$. Then, γ is successful and*

- (1) *if every action γ_t is optimal at $p_t(\gamma)$ then strategy γ is optimal;*
- (2) *for every $\varepsilon > 0$ and $\delta \in [0, \alpha\varepsilon]$, if every action $\gamma_t \in \{L, R\}$ is δ -optimal at $p_t(\gamma)$ and every action $\gamma_t = 0$ is optimal at $p_t(\gamma)$, then strategy γ is ε -optimal.*

Proof. Since γ never waits infinitely long, it can be described as follows: First, wait until period w_1 , then search location $a_1 \in \{L, R\}$, then wait until period w_2 , then search location a_2 , and so on. For $k = 1, 2, 3, \dots$ define

$$z_k := \begin{cases} p_{w_k}(\gamma), & \text{if } a_k = L \\ 1 - p_{w_k}(\gamma), & \text{if } a_k = R. \end{cases}$$

From now on, we will write p_t instead of $p_t(\gamma)$. Hence, z_k denotes the probability that the object will be found by action a_k at period w_k , if the object has not been found yet, and strategy γ is being implemented. Note that, by our assumption in the lemma, $z_k \geq \alpha$ for all k .

We now show that γ is successful. The probability that γ never finds the object is $\prod_{k \geq 1} (1 - z_k)$. Since $z_k \geq \alpha$ for all k , we have that $1 - z_k \leq 1 - \alpha < 1$. Hence, $\prod_{k \geq 1} (1 - z_k) = 0$, yielding that γ is successful.

We now prove (1) and (2) by showing that for every $\varepsilon \geq 0$, if $\delta \in [0, \alpha\varepsilon]$, every action $\gamma_t \in \{L, R\}$ is δ -optimal at $p_t(\gamma)$, and every action $\gamma_t = 0$ is optimal at $p_t(\gamma)$,

then strategy γ is ε -optimal. Part (1) would then follow by taking $\varepsilon = 0$, and part (2) would follow by taking $\varepsilon > 0$.

Choose $\varepsilon \geq 0$ and $\delta \in [0, \alpha\varepsilon]$. Suppose now that at every period t every action $\gamma_t \in \{L, R\}$ is δ -optimal for $p_t(\gamma)$, and every action $\gamma_t = 0$ is optimal at $p_t(\gamma)$. We will show that γ is an ε -optimal strategy, i.e. $C(\bar{p}, \gamma) \leq V(\bar{p}) + \varepsilon$, where \bar{p} is the initial probability that the object is at L .

For every period t , let γ^t be the continuation strategy from period t onwards. Then,

$$C(\bar{p}, \gamma) = z_1 \cdot 1 + (1 - z_1) [1 + C(p_{w_1+1}, \gamma^{w_1+1})],$$

since the object will be found at period w_1 with probability z_1 , while with probability $1 - z_1$ the continuation strategy γ^{w_1+1} would be played after period w_1 . On the other hand,

$$V(\bar{p}) = V(p_{w_1}) \geq V(p_{w_1}, a_1) - \delta = z_1 \cdot 1 + (1 - z_1) [1 + V(p_{w_1+1})] - \delta.$$

Here, the first equality follows from the fact that waiting until period w_1 is a sequence of optimal actions, and the inequality follows from the assumption that a_1 is a δ -optimal action at period w_1 . The last equality follows from (1) and (2). Therefore,

$$V(\bar{p}) - C(\bar{p}, \gamma) \geq (1 - z_1) [V(p_{w_1+1}) - C(p_{w_1+1}, \gamma^{w_1+1})] - \delta,$$

and by induction it would follow that

$$\begin{aligned} V(\bar{p}) - C(\bar{p}, \gamma) &\geq (1 - z_1) \cdot \dots \cdot (1 - z_k) [V(p_{w_k+1}) - C(p_{w_k+1}, \gamma^{w_k+1})] \\ &\quad - \delta [1 + (1 - z_1) + \dots + (1 - z_1) \cdots (1 - z_{k-1})] \end{aligned} \quad (7)$$

for $k = 1, 2, 3, \dots$

We will show that $V(\bar{p}) - C(\bar{p}, \gamma) \geq -\varepsilon$, which would imply that γ is an ε -optimal strategy. We already know that $\prod_{k \geq 1} (1 - z_k) = 0$. Hence, it is sufficient to show that (i) $V(p_{w_k+1}) - C(p_{w_k+1}, \gamma^{w_k+1})$ is uniformly bounded, and

$$(ii) \delta [1 + (1 - z_1) + \dots + (1 - z_1) \cdots (1 - z_{k-1})] \leq \varepsilon$$

for all k .

(i) Since, by Lemma 3.1 (i), $1 \leq V(p_{w_k+1}) \leq 2$, it remains to prove that $C(p_{w_k+1}, \gamma^{w_k+1})$ is uniformly bounded.

As $z_k \geq \alpha$ for all k , it holds that

$$\begin{aligned} C(p_{w_k+1}, \gamma^{w_k+1}) &\leq \alpha \cdot 1 + (1 - \alpha) \cdot \alpha \cdot 2 + (1 - \alpha)^2 \cdot \alpha \cdot 3 + \dots \\ &= \alpha(1 + (1 - \alpha) + (1 - \alpha)^2 + \dots)^2 \\ &= \alpha \frac{1}{\alpha^2} = \frac{1}{\alpha}, \end{aligned}$$

and hence $C(p_{w_k+1}, \gamma^{w_k+1})$ is uniformly bounded by $\frac{1}{\alpha}$.

(ii) For every k ,

$$\begin{aligned} \delta [1 + (1 - z_1) + \dots + (1 - z_1) \cdots (1 - z_{k-1})] &\leq \alpha\varepsilon [1 + (1 - \alpha) + \dots + (1 - \alpha)^{k-1}] \\ &\leq \alpha\varepsilon \frac{1}{\alpha} = \varepsilon. \end{aligned}$$

This completes the proof of the lemma. ■

The next result relates the transition probabilities, q and r , to the values of π^* and $\bar{\pi}$, which will be useful for the proof of the main theorem.

Lemma 5.2 *For any transition probabilities $q, r \in [0, 1]$ with $q + r > 0$, it holds that (i) $q < r \iff \bar{\pi} < \pi^*$, (ii) $q > r \iff \bar{\pi} > \pi^*$ and (iii) $q = r \iff \bar{\pi} = \pi^*$.*

Proof. It is sufficient to show that the following three implications are valid: (i) $q < r \implies \bar{\pi} < \pi^*$, (ii) $q > r \implies \bar{\pi} > \pi^*$, (iii) $q = r \implies \bar{\pi} = \pi^*$.

Proof of (i). Suppose $q < r$. In view of (5) and (6), $\bar{\pi} < \pi^*$ is equivalent to

$$qV(r) < rV(1 - q). \quad (8)$$

Thus, we have to show that (8) holds. If $V(r) < V(1 - q)$, then (8) is obvious. So suppose $V(r) > V(1 - q)$. This means that $r \neq 1 - q$, which together with the assumption $q < r$ leaves us only two cases: $q + r \in (0, 1)$ (the non-oscillating case) and $q + r \in (1, 2)$ (the oscillating case).

Assume $q + r \in (0, 1)$ (the non-oscillating case). By (6) we have $r < \pi^* < 1 - q$. Since $A^n r$ is monotonically converging to π^* , we have for all $n \in \mathbb{N}$, $V(r, L) > V(A^n r, L) > V(A^{n+1}, L)$ and $\lim_{n \rightarrow \infty} V(A^n r, L) = V(\pi^*, L)$, as well as $V(r, R) < V(A^n r, R)$. Hence by property (ii) of Lemma 3.1 we obtain

$$V(r) = \min\{V(\pi^*, L), V(r, R)\}.$$

Similarly,

$$V(1 - q) = \min\{V(1 - q, L), V(\pi^*, R)\}$$

Since $V(\pi^*, R) > V(r, R)$, by our assumption $V(r) > V(1 - q)$, inequality $V(1 - q, L) < V(r, R)$ follows. Hence, by (1) and (2),

$$1 + qV(r) = V(1 - q, L) < V(r, R) = 1 + rV(1 - q)$$

yielding (8).

Assume now $q + r \in (1, 2)$ (the oscillating case). By (6) we have $1 - q < \pi^* < r$. Due to the facts that $A(1 - q) > (1 - q)$ and $A(1 - q) > A^n(1 - q)$ for all $n > 1$ we can conclude that $V(A(1 - q), L) < V(1 - q, L)$ and $V(A(1 - q), L) < V(A^n(1 - q), L)$ for all $n > 1$.

Also by $(1 - q) < A^n(1 - q)$ for all $n \in \mathbb{N}$ we have $V(1 - q, R) < V(A^n(1 - q), R)$ for all $n \in \mathbb{N}$. Therefore, by property (ii) of Lemma 3.1, $V(1 - q) = \min\{V(A(1 - q), L), V(1 - q, R)\}$ follows.

Similarly to the arguments above, $V(r) = \min\{V(r, L), V(Ar, R)\}$. Since, by (4), $A(1 - q) < r$, we obtain $V(A(1 - q), L) > V(r, L)$. Hence, by $V(r) > V(1 - q)$, we must have $V(1 - q, R) < V(r, L)$ yielding $V(1 - q) = V(1 - q, R)$. Therefore,

$$V(1 - q) = V(1 - q, R) = 1 + (1 - q)V(1 - q)$$

Now, we need to analyze two cases: (a) $q \neq 0$ and (b) $q = 0$.

For (a), $V(1 - q) = \frac{1}{q}$. Therefore, in view of property (v) of Lemma 3.1, we may conclude that $qV(r) \leq \frac{q}{r} < \frac{r}{q} = rV(1 - q)$ which implies (8).

For (b), $qV(r) = 0 < rV(1 - q)$, hence (8).

Proof of (ii). Similar to the proof of (i) above.

Proof of (iii). Assume that $q = r$. Then, as transition probabilities q and r are equal, the problem is symmetric, and $V(p) = V(1 - p)$ for all $p \in [0, 1]$. Hence, by (5) and (6), we obtain $\bar{\pi} = \frac{1}{2}$ and $\pi^* = \frac{1}{2}$, hence $\bar{\pi} = \pi^*$ follows. ■

The reader is referred to the appendix for a graphical representation of the functions $V(p, L)$ and $V(p, R)$ in each of the cases $q < r$, $q = r$ and $q > r$.

6 Existence of Thresholds

In this section we will show that we can always find thresholds π_1 and π_2 such that the induced strategy γ is (ε) -optimal and successful. We also show that these thresholds are unique if $q = r$ or $q + r \geq 1$. Consequently, we prove for these cases the uniqueness of the optimal strategy, up to the choices at the two thresholds. The exact computation of the thresholds follows in Section 7.

In view of Lemma 5.1, it is sufficient to prove the following two lemmas.

Lemma 6.1 *Take a configuration (\bar{p}, q, r) of initial and transition probabilities with $q \leq r$. Then, there are thresholds $\pi_1, \pi_2 \in [0, 1]$ with $\pi_1 \leq \pi_2$, $\pi_1 < 1$ and $\pi_2 > 0$ such that: searching R is an optimal action for every $p \in [0, \pi_1]$, waiting is an optimal action for every $p \in (\pi_1, \pi_2)$, and searching L is an optimal action for every $p \in [\pi_2, 1]$. Here, the optimal actions are unique up to the choices at π_1 and π_2 . As usual, p denotes the probability for the object being located at L .*

Moreover, if $q = r$ or $q + r \geq 1$, then the induced strategy γ never waits infinitely long.

Also, if $q + r > 0$, then $\pi_2 = \pi^$.*

We will prove this lemma later in Sections 6.1–6.5.

Lemma 6.2 *Take a configuration (\bar{p}, q, r) of initial and transition probabilities such that $0 < q + r < 1$ (i.e., the non-oscillating case) and $q < r$. Then, there is a threshold $\pi_1 \in [0, \pi^*]$ satisfying $\pi_1 < 1$ with the following property: For every $\delta > 0$ there is a $\tau > 0$ such that searching R is an optimal action for every $p \in [0, \pi_1]$, waiting is an optimal action for every $p \in (\pi_1, \pi^* - \tau)$, and searching L is a δ -optimal action for every $p \in [\pi^* - \tau, 1]$. Moreover, the induced strategy γ never waits infinitely long.*

We will prove this lemma in Section 6.6.

We will now prove Lemma 6.1. We distinguish five cases.

6.1 Absorbing case: $q + r = 0$

The fact that $q = r = 0$ implies directly that the object does not move. Hence, waiting makes no sense. By searching a location at period 1, one either finds the object immediately, or one will be sure to find the object next period at the other location. Therefore, only L is optimal if $p > \frac{1}{2}$, only R is optimal if $p < \frac{1}{2}$, and both are optimal at $p = \frac{1}{2}$. Accordingly, $\pi_1 = \pi_2 = \bar{\pi} = \frac{1}{2}$. Note that the optimal actions are unique up to the choice at $\frac{1}{2}$, and moreover the induced strategy never waits.

6.2 Non-oscillating case: $0 < q + r < 1$

Since $q \leq r$, by Lemma 5.2 we have that $\bar{\pi} \leq \pi^*$. Define $\pi_2 = \pi^*$ and let $\pi_1 \in [0, \bar{\pi}]$ be the unique p for which $V(p, R) = V(\pi^*, L)$. Note that π_1 is well-defined for the following reasons: First, $V(0, R) = 1 \leq V(\pi^*, L)$ due to the Lemma 3.1, (i). Also, $V(\bar{\pi}, R) = V(\bar{\pi}, L) \geq V(\pi^*, L)$ due to the definition of $\bar{\pi}$, monotonicity of $V(p, L)$, and the fact that $\pi^* \geq \bar{\pi}$. Hence, there exists a $p \in [0, \bar{\pi}]$ with $V(p, R) = V(\pi^*, L)$. Moreover, since $V(p, R)$ is strictly monotone, there can only be one such p .

Note that $\pi_2 > 0$ as $\pi_2 = \pi^* = \frac{r}{q+r}$, $r \geq q$ and $q + r > 0$. Also, as $\bar{\pi} < 1$ we have that $\pi_1 < 1$.

On the interval $[0, \pi_1]$: It holds that $V(p, R) < V(p, L)$ since $\pi_1 \leq \bar{\pi}$. Since $p < A^n p$ for all n , we have that $V(p, R) < V(A^n p, R)$ for all n . Moreover, since $A^n p < \pi^*$, it holds that $V(p, R) < V(\pi_1, R) = V(\pi^*, L) \leq V(A^n p, L)$. Therefore, by Lemma 3.1, (ii), $V(p) = V(p, R)$, and only R is optimal at $[0, \pi_1]$.

It similarly follows that R is optimal at π_1 (and waiting as well).

On the interval (π_1, π_2) : We have that $V(p, R) > V(\pi_1, R) = V(\pi^*, L) \geq V(\pi^*)$, and $V(p, L) > V(\pi_2, L) = V(\pi^*, L) \geq V(\pi^*)$. Since, by Lemma 3.1, (iv), $V(p) \leq V(\pi^*)$, it follows that $V(p) = V(p, 0)$, and therefore only waiting is optimal at (π_1, π_2) .

On the interval $(\pi_2, 1]$: We know that $V(p, L) < V(p, R)$. Since $A^n p < p$ for all n , we have that $V(p, L) < V(A^n p, L)$ for all n . Moreover, since $A^n p > \pi^*$, it holds that $V(p, L) < V(A^n p, L) \leq V(A^n p, R)$. Therefore, by Lemma 3.1, (ii), $V(p) = V(p, L)$, and only L is optimal at $(\pi_2, 1]$.

It similarly follows that L is optimal at π_2 (and waiting as well).

So, the optimal actions are unique up to the choice at π_1 and π_2 .

Note that if $q = r$, then $\pi^* = \bar{\pi}$ by Lemma 5.2, hence $\pi_1 = \pi_2 = \pi^*$. Therefore, the induced strategy never waits.

6.3 State independent transitions case: $q + r = 1$

The proof for this case is identical to the previous case.

6.4 Oscillating case: $1 < q + r < 2$

Since $q \leq r$, by Lemma 5.2 we have that $\bar{\pi} \leq \pi^*$. Define $\pi_2 = \pi^*$ and let $\pi_1 \in [0, \bar{\pi}]$ be the unique p for which $V(p, R) = V(Ap, L)$. Note that π_1 is well-defined for the following reasons: First, $V(0, R) = 1 \leq V(A0, L)$ due to the Lemma 3.1, (i). Also,

since $\pi^* \geq \bar{\pi}$, we have $V(\bar{\pi}, R) = V(\bar{\pi}, L) \geq V(\pi^*, L) \geq V(A\bar{\pi}, L)$, where the last inequality follows from the fact that $A\bar{\pi} \geq \pi^*$. Hence, there exists a $p \in [0, \bar{\pi}]$ with $V(p, R) = V(Ap, L)$. Now, suppose that more than one $p \in [0, 1]$ would exist with $V(p, R) = V(Ap, L)$. Since both sides are linear functions in p , it must be the case that $V(p, R) = V(Ap, L)$ for all $p \in [0, 1]$. In particular, $1 = V(0, R) = V(A0, L)$, which by Lemma 3.1, (i) implies that $A0 = 1$, and hence $r = 1$. Also, $V(\bar{\pi}, L) = V(\bar{\pi}, R) = V(A\bar{\pi}, L)$, which implies that $\bar{\pi} = A\bar{\pi}$, and hence $\bar{\pi} = \pi^*$. It follows from Lemma 5.2 that $q = r$, and hence $q = r = 1$, which contradicts $q + r < 2$. So, there can only be one p with $V(p, R) = V(Ap, L)$.

Note that $\pi_2 > 0$ as $\pi_2 = \pi^* = \frac{r}{q+r}$, $r \geq q$ and $q + r > 1$. Also, as $\bar{\pi} < 1$ we have that $\pi_1 < 1$.

It also follows that

$$V(p, R) < V(Ap, L) \text{ for all } p \in [0, \pi_1), \text{ and} \quad (9)$$

$$V(p, R) > V(Ap, L) \text{ for all } p \in (\pi_1, 1]. \quad (10)$$

On the interval $[0, \pi_1)$ it holds that $V(p, R) < V(p, L)$ since $\pi_1 \leq \bar{\pi}$. Since $A^n p > p$ for all n , we have that $V(p, R) < V(A^n p, R)$ for all n . Moreover, by (9), $V(p, R) < V(Ap, L) \leq V(A^n p, L)$, where the latter follows from the fact that $A^n p \leq Ap$ for all n . Therefore, by Lemma 3.1, (ii), $V(p) = V(p, R)$, and therefore only R is optimal at $[0, \pi_1)$.

It similarly follows that R is optimal at π_1 (and waiting as well).

On the interval (π_1, π_2) , we have that $V(Ap, L) < V(p, L)$ since $Ap > p$. Moreover, $V(Ap, L) < V(p, R)$ by (10). In view of $V(p, 0) = V(Ap) \leq V(Ap, L)$, it follows that only waiting is an optimal action at (π_1, π_2) .

On the interval $(\pi_2, 1]$, we know that $V(p, L) < V(p, R)$. Since $p > A^n p$ for all n , we have that $V(p, L) < V(A^n p, L)$ for all n . Since $\pi^* \geq \bar{\pi}$, we have that $V(\pi^*, L) \leq V(\pi^*, R) = V(A\pi^*, R)$. Also, $V(1, L) = 1 < V(A1, R)$ as $A1 = 1 - q > 0$. Since $V(p, L)$ and $V(Ap, R)$ are linear in p , it follows that $V(p, L) < V(Ap, R)$ on $(\pi_2, 1]$. Moreover, $V(Ap, R) \leq V(A^n p, R)$ since $A^n p \geq Ap$ for all n . Therefore, $V(p, L) < V(A^n p, R)$ for all n . Together with $V(p, L) < V(A^n p, L)$ for all n , it follows from Lemma 3.1, (ii), that $V(p) = V(p, L)$, and only L is optimal at $(\pi_2, 1]$.

It similarly follows that L is optimal at π_2 (and waiting as well).

So, the optimal actions are unique up to the choice at π_1 and π_2 , and the induced strategy waits at most once.

6.5 Switching case: $q + r = 2$

In this case, $q = r = 1$, so the object will surely move to the other location if it is not found. By searching a location at period 1, one either finds the object immediately, or one will be sure to find the object next period at the same location. Therefore, only L is optimal if $p > \frac{1}{2}$, only R is optimal if $p < \frac{1}{2}$, and both are optimal at $p = \frac{1}{2}$. Accordingly, $\pi_1 = \pi_2 = \bar{\pi} = \frac{1}{2}$. Note that the optimal actions are unique up to the choice at $\frac{1}{2}$, and moreover the induced strategy never waits.

6.6 Near Optimality in the Non-Oscillating Case: $0 < q + r < 1$ and $q < r$

We now prove Lemma 6.2. Choose π_1 as in Lemma 6.1. Take a $\delta > 0$, and choose $\tau > 0$ small enough such that $\pi^* - \tau \geq \bar{\pi}$ and $V(\pi^* - \tau, L) - V(\pi^*, L) \leq \delta$. Such a τ exists due to Lemma 5.2 (i), and the continuity of $V(p, L)$. On the interval $[\pi^* - \tau, \pi^*]$ we have that (1) $V(p, R) \geq V(p, L)$, (2) $A^n p$ converges to π^* in a monotonically increasing way, and (3) $V(p, L)$ is decreasing in p . Therefore, by Lemma 3.1 (ii), $V(p) = V(\pi^*, L)$ for every $p \in [\pi^* - \tau, \pi^*]$. Now, take some $p \in [\pi^* - \tau, \pi^*]$. Then,

$$V(p, L) - V(p) = V(p, L) - V(\pi^*, L) \leq V(\pi^* - \tau, L) - V(\pi^*, L) \leq \delta,$$

which means that searching L is a δ -optimal action on $[\pi^* - \tau, \pi^*]$. As on the interval $[\pi^*, 1]$, searching L is an optimal action (cf. Section 6.2), we obtain that searching L is an optimal action on the whole $[\pi^* - \tau, 1]$. It is clear that the strategy induced by π_1 and $\pi^* - \tau$ never waits infinitely long.

7 Computation of Thresholds

In this section we derive exact formulas for the thresholds π_1 and π_2 in Lemmas 6.1 and 6.2. In Sections 7.1 we deal with Lemma 6.1, and also show that $\pi_1 \leq \frac{1}{2}$ and $\pi_2 \geq \frac{1}{2}$, which will imply that the induced strategies only prescribe to search a location if the probability that this location contains the object is at least $\frac{1}{2}$. In Section 7.2 we deal with the computation for Lemma 6.2.

7.1 Computation for Lemma 6.1

Note that for the cases $q + r = 0$ and $q + r = 2$ we have already shown that $\pi_1 = \pi_2 = \frac{1}{2}$. So, it remains to analyze the cases where $0 < q + r < 1$, $q + r = 1$, and $1 < q + r < 2$. Recall our assumption that $q \leq r$. By Lemma 5.2 we know that $\bar{\pi} \leq \pi^*$. Recall that $\pi_2 = \pi^* = \frac{r}{q+r} \geq \frac{1}{2}$. So, we only need to compute π_1 and show that $\pi_1 \leq \frac{1}{2}$.

7.1.1 Non-oscillating case: $0 < q + r < 1$

Since $q + r < 1$, we have $\pi^* > r$, and $1 - \pi^* = \frac{q}{q+r} > q$, which implies $\pi^* < 1 - q$. So,

$$r < \pi^* < 1 - q.$$

We distinguish two cases here: (i) $r \leq \pi_1$ and (ii) $r > \pi_1$. We will show that case (i) corresponds to the case in Table 1 where $(q+1)r^2 + q^2r - q \leq 0$, and that case (ii) corresponds to the case where $(q+1)r^2 + q^2r - q > 0$.

Case (i) $r \leq \pi_1$. By Lemma 6.1 we know that

$$\begin{aligned} V(r) &= V(r, R), \text{ and} \\ V(1 - q) &= V(1 - q, L) \text{ since } 1 - q > \pi^* = \pi_2. \end{aligned}$$

From equations (1) and (2), it follows

$$V(r) = 1 + rV(1 - q) \text{ and } V(1 - q) = 1 + qV(r),$$

which implies that

$$V(r) = \frac{1 + r}{1 - rq} \text{ and } V(1 - q) = \frac{1 + q}{1 - rq}.$$

Then, by equations (1) and (2) we can write

$$V(p, R) = 1 + p\frac{1 + q}{1 - rq} \text{ and } V(p, L) = 1 + (1 - p)\frac{1 + r}{1 - rq}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(\pi^*, L)$. Since

$$V(\pi_1, R) = 1 + \pi_1\frac{1 + q}{1 - rq} \text{ and } V(\pi^*, L) = 1 + (1 - \frac{r}{q + r})\frac{1 + r}{1 - rq},$$

we get

$$\pi_1 = \frac{q(1 + r)}{(q + r)(1 + q)}.$$

An easy calculation shows that $r \leq \pi_1$ if and only if $(q + 1)r^2 + q^2r - q \leq 0$. Also, it may be verified that $q \leq r$ implies $\pi_1 \leq \frac{1}{2}$.

Case (ii) $r > \pi_1$. Since $r < \pi^*$, we know from Lemma 6.1 that

$$V(r) = V(r, 0) = V(Ar).$$

Since $Ar \in (\pi_1, \pi^*)$, we obtain by induction that

$$V(r) = V(A^n r)$$

for all n , and hence

$$V(r) = \lim_{n \rightarrow \infty} V(A^n r) = V(\pi^*).$$

As $\pi^* \geq \bar{\pi}$, we have by Lemma 3.1, (iii) that $V(\pi^*) = V(\pi^*, L)$. Hence,

$$V(r) = V(\pi^*, L).$$

Since $1 - q > \pi^* = \pi_2$, Lemma 6.1 yields $V(1 - q) = V(1 - q, L)$.

Combining these insights with equations (1) and (2) we obtain

$$V(1 - q) = 1 + qV(r) \text{ and } V(r) = 1 + (1 - \pi^*)V(r).$$

From these two equations we get

$$V(r) = \frac{q + r}{r} \text{ and } V(1 - q) = 1 + q\frac{q + r}{r}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(\pi^*, L)$. Since

$$V(\pi_1, R) = 1 + \pi_1 V(1 - q) \text{ and } V(\pi^*, L) = V(r) = \frac{q + r}{r},$$

we obtain that

$$\pi_1 = \frac{q}{(q + r)(1 + q) - q}.$$

An easy calculation shows that $r > \pi_1$ if and only if $(q + 1)r^2 + q^2r - q > 0$.

We will now show that $\pi_1 \leq \frac{1}{2}$. If $r \leq \frac{1}{2}$, then $\pi_1 < r$ yields $\pi_1 \leq \frac{1}{2}$. Suppose now that $r > \frac{1}{2}$. Let $\phi(a, b) := a^2 + a(b - 2) + b$ for all real numbers a and b . Then, the inequality $\pi_1 \leq \frac{1}{2}$ is equivalent with $\phi(q, r) \geq 0$. Notice that, given any b , the parabola $\phi(a, b)$ is minimal at $a^* = 1 - \frac{b}{2}$. By using that

$$q < 1 - r \leq 1 - \frac{r}{2},$$

we obtain

$$\phi(q, r) > \phi(1 - r, r) = -1 + 2r \geq 0,$$

since $r > \frac{1}{2}$. Hence, $\pi_1 \leq \frac{1}{2}$.

7.1.2 State independent transitions case: $q + r = 1$

In this case, $\pi_2 = \pi^* = r = 1 - q$. Since $q \leq r$, we know that $q \leq \frac{1}{2}$. We show that $\pi_1 = q$, which will imply that $\pi_1 \leq \frac{1}{2}$. By copying the arguments above, we obtain the following:

Case (i) $r \leq \pi_1$. We find that

$$\pi_1 = \frac{q(1 + r)}{(q + r)(1 + q)} = \frac{q(1 + r)}{(1 + q)}.$$

We note that $r \leq \pi_1$ and $q \leq r$ yield $q = r$, and hence $\pi_1 = q$.

Case (ii) $r > \pi_1$. We get

$$\pi_1 = \frac{q}{(q + r)(1 + q) - q} = q.$$

7.1.3 Oscillating case: $1 < q + r < 2$

Since $q + r > 1$, we have $\pi^* < r$, and $1 - \pi^* = \frac{q}{q + r} < q$, which implies $\pi^* > 1 - q$. So,

$$1 - q < \pi^* < r.$$

We distinguish two cases here: (i) $1 - q \leq \pi_1$ and (ii) $1 - q > \pi_1$. We will show that case (i) corresponds to the case in Table 1 where $(1 - q + q^2)r - 2q^2 + q^3 \leq 0$, while case (ii) corresponds to the case where $(1 - q + q^2)r - 2q^2 + q^3 > 0$.

Case (i) $1 - q \leq \pi_1$. Since $\pi^* = \pi_2$, we know by Lemma 6.1 that

$$V(1-q) = V(1-q, R) \text{ and } V(r) = V(r, L).$$

From equations (1) and (2), it follows

$$V(1-q) = 1 + (1-q)V(1-q) \text{ and } V(r) = 1 + (1-r)V(r).$$

Hence

$$V(1-q) = \frac{1}{q} \text{ and } V(r) = \frac{1}{r}.$$

Then, by equations (1) and (2) we can write

$$V(p, R) = 1 + p\frac{1}{q} \text{ and } V(p, L) = 1 + (1-p)\frac{1}{r}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(A\pi_1, L)$. Since $A\pi_1 = (1-q-r)\pi_1 + r$, we have

$$V(\pi_1, R) = 1 + \pi_1\frac{1}{q} \text{ and } V(A\pi_1, L) = 1 + (1 - (1-q-r)\pi_1 - r)\frac{1}{r}.$$

This yields

$$\pi_1 = \frac{q(1-r)}{(q+r)(1-q)}.$$

An easy calculation shows that $1-q \leq \pi_1$ if and only if $(1-q+q^2)r - 2q^2 + q^3 \leq 0$. It can also be verified that $q \leq r$ implies $\pi_1 \leq \frac{1}{2}$.

Case (ii) $1-q > \pi_1$. Since $\pi^* = \pi_2$, we know by Lemma 6.1 that

$$V(1-q) = V(1-q, 0) \text{ and } V(r) = V(r, L).$$

As $V(1-q, 0) = V(A(1-q))$ and $A(1-q) > \pi^* = \pi_2$, we obtain $V(1-q, 0) = V(A(1-q), L)$. Since $A(1-q) = (1-q-r)(1-q) + r$, it follows from equations (1) and (2) that

$$V(1-q) = 1 + (1 - (1-q-r)(1-q) - r)V(r) \text{ and } V(r) = 1 + (1-r)V(r).$$

Hence

$$V(r) = \frac{1}{r} \text{ and } V(1-q) = \frac{q + (q+r)(1-q)}{r}.$$

Then, by equations (1) and (2) we can write

$$V(p, R) = 1 + p\frac{q + (q+r)(1-q)}{r} \text{ and } V(p, L) = 1 + (1-p)\frac{1}{r}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(A\pi_1, L)$. Since $A\pi_1 = (1-q-r)\pi_1 + r$, we have

$$\begin{aligned} V(\pi_1, R) &= 1 + \pi_1\frac{q + (q+r)(1-q)}{r} \text{ and} \\ V(A\pi_1, L) &= 1 + (1 - (1-q-r)\pi_1 - r)\frac{1}{r}. \end{aligned}$$

This yields

$$\pi_1 = \frac{1-r}{(q+r)(1-q) + 1-r}.$$

An easy calculation shows that $1-q > \pi_1$ if and only if $(1-q+q^2)r - 2q^2 + q^3 > 0$.

We will now show that $\pi_1 \leq \frac{1}{2}$. As $q+r > 1$ and $q \leq r$, we have $r > \frac{1}{2}$. Let $\phi(a, b) := -a^2 + a(1-b) + (2b-1)$ for all real numbers a and b . Then, the inequality $\pi_1 \leq \frac{1}{2}$ is equivalent with $\phi(q, r) \geq 0$. Notice that, given any b , the parabola $\phi(a, b)$ is maximal at $a^* = \frac{1-b}{2}$. By using that

$$\frac{1-r}{2} < 1-r < q \leq r$$

we obtain

$$\phi(q, r) > \phi(r, r) = -2r^2 + 3r - 1 = (1-r)(2r-1) \geq 0,$$

since $r > \frac{1}{2}$. Hence, $\pi_1 \leq \frac{1}{2}$.

7.2 Computation for Lemma 6.2

We now compute the threshold $\pi^* - \tau$ in Lemma 6.2. Recall that $0 < q+r < 1$ and $q < r$. Take a $\delta > 0$, and choose $\tau := \min\{\pi^* - \frac{1}{2}, \frac{\delta}{2}\}$. Note that $\tau > 0$ and $\pi^* - \tau \geq \frac{1}{2}$. We know, from Section 7.1.1, that $\pi_1 \leq \frac{1}{2}$, and hence $\pi^* - \tau \geq \frac{1}{2} \geq \pi_1$. Therefore, on the interval $[\pi^* - \tau, \pi^*]$ we have that (1) $V(p, R) \geq V(\pi^*, L)$, (2) $V(p, L) \geq V(\pi^*, L)$, and (3) $A^n p$ converges to π^* in a monotonically increasing way. Hence, by Lemma 3.1 (ii), $V(p) = V(\pi^*, L)$ for every $p \in [\pi^* - \tau, \pi^*]$. Now, take some $p \in [\pi^* - \tau, \pi^*]$. In order to prove that searching L is a δ -optimal action, it is sufficient to show that

$$V(p, L) \leq V(p) + \delta = V(\pi^*, L) + \delta.$$

We know, by (1), that

$$V(p, L) = 1 + (1-p)V(r)$$

and

$$V(\pi^*, L) = 1 + (1-\pi^*)V(r).$$

Hence $V(p, L) - V(\pi^*, L) = (\pi^* - p)V(r)$. But $(\pi^* - p) \leq \frac{\delta}{2}$ by the choice of p . Moreover, $V(r) \leq 2$ by Lemma 3.1, (i). Therefore,

$$V(p, L) - V(\pi^*, L) = (\pi^* - p)V(r) \leq \delta.$$

8 Relation to the Model without Waiting

In this section we compare the optimal strategy we determined in the previous sections with the optimal strategy that Pollock (1970) found for the model without waiting. We thus evaluate the precise consequence of introducing the option to wait. Recall that the optimal strategy in Pollock (1970) is given by only one threshold $\tilde{\pi}$, meaning

that the agent searches R if $p < \tilde{\pi}$ and searches L if $p > \tilde{\pi}$, and is indifferent if $p = \tilde{\pi}$. We first illustrate, by means of an example from Ross (1983), that in the model without waiting the agent may search a location that contains the object with probability less than one half (i.e. $\tilde{\pi} \neq \frac{1}{2}$). Afterwards, we show that Pollock's threshold $\tilde{\pi}$ is always in between our thresholds π_1 and π_2 . Consequently, if the agent searches a location in the model with waiting, he would search the same location in the model without waiting. Finally, we compare the values of the two models.

8.1 Ross' Example Revisited

An important implication of our model is that it is never optimal to search the location with lower current probability of containing the object (See Theorem 2.2). In the model without the option to wait, as Ross (1983) has shown, it is possible that searching the location with the lower probability of containing the object is optimal. In fact, this is a striking result of the model without waiting. This happens if searching the location with the lower probability of containing the object serves as an investment for the future, i.e. decreases the uncertainty about the location of the object in the future. The following example by Ross (1983) shows this.

Example (Ross, 1983) $\bar{p} = 0.45$ and $1 - \bar{p} = 0.55$, $q = \frac{1}{2}$ and $r = 1$.

Consider first the model without waiting. Although the initial probability of location R containing the object is higher, it is optimal to search L since searching L gives complete certainty about the location of the object at $t = 2$ in case it is not found at $t = 1$. One can see that, in contrast to searching L , searching R at $t = 1$ leads to complete uncertainty about the location of the object at $t = 2$, in case it is not found at $t = 1$. As a result, the expected cost induced by searching L at $t = 1$ and acting optimally afterwards (i.e., search L again) is 1.55 and the expected cost induced by searching R at $t = 1$ and acting optimally afterwards (i.e., search L , and if the object is not found then search L once more) is 1.675. In fact, it follows from Pollock's calculations that $\tilde{\pi} = 0.4 < \frac{1}{2}$.

Let's now introduce the option to wait. Since $1 < q + r < 2$ and $(1 - q + q^2)r - 2q^2 + q^3 > 0$ in this example, we see from Table 1 that the thresholds are $\pi_1 = 0$ and $\pi_2 = \frac{2}{3}$. Notice the large difference with the model without waiting. Hence, at period 1 the agent would wait instead of searching L . After waiting, the probability that the object is located at L is $(0.45)(0.5) + (0.55)1 = 0.775 > \pi_2$. Hence, the agent would search L at period 2. If at period 2, he does not find the object, he will surely find it in period 3 by searching L . Accordingly, the expected cost of finding the object is $(0.775) \cdot 1 + (0.225) \cdot 2 = 1.225$, which is less than the expected searching cost without the option to wait (i.e., 1.55).

8.2 Comparison of Thresholds

We will now explicitly compare our thresholds π_1 and π_2 with Pollock's threshold $\tilde{\pi}$ for the model without waiting. We show that, for any configuration of initial and transition probabilities, we have $\pi_1 \leq \tilde{\pi} \leq \pi_2$. The computations that follow largely resemble Pollock's analysis.

For the model without waiting, let $\tilde{V}(p)$ denote the value function, while $\tilde{V}(p, L)$ and $\tilde{V}(p, R)$ the expected costs induced by searching L and R , respectively, given that the agent behaves optimally afterwards. Then, similarly to equalities (1) and (2), we obtain

$$\tilde{V}(p, L) = p \cdot 1 + (1 - p) \left[1 + \tilde{V}(r) \right] = 1 + (1 - p)\tilde{V}(r) \quad (11)$$

and

$$\tilde{V}(p, R) = (1 - p) \cdot 1 + p \left[1 + \tilde{V}(1 - q) \right] = 1 + p\tilde{V}(1 - q). \quad (12)$$

Then, the unique intersection point of $\tilde{V}(p, L)$ and $\tilde{V}(p, R)$ is exactly the threshold $\tilde{\pi}$, which is given by

$$\tilde{\pi} = \frac{\tilde{V}(r)}{\tilde{V}(1 - q) + \tilde{V}(r)}. \quad (13)$$

Thus, in the model without waiting, only action R is optimal (i.e. $\tilde{V}(p) = \tilde{V}(p, R) < \tilde{V}(p, L)$) when $p < \tilde{\pi}$, both actions L and R are optimal (i.e. $\tilde{V}(p) = \tilde{V}(p, L) = \tilde{V}(p, R)$) when $p = \tilde{\pi}$ and only action L is optimal (i.e. $\tilde{V}(p) = \tilde{V}(p, L) < \tilde{V}(p, R)$) when $p > \tilde{\pi}$.

Now we show that, for any configuration of initial and transition probabilities, we have $\pi_1 \leq \tilde{\pi} \leq \pi_2$. Again, we assume that $q \leq r$. We distinguish the following cases.

Case 1. $q + r = 0$. Recall from Section 6.1 that $\pi_1 = \pi_2 = \frac{1}{2}$. Obviously, $q + r = 0$ yields $q = r = 0$. By $\tilde{V}(r) = \tilde{V}(0) = \tilde{V}(0, R) = 1$ and $\tilde{V}(1 - q) = \tilde{V}(1) = \tilde{V}(1, L) = 1$, we obtain $\tilde{\pi} = \frac{1}{2}$.

Case 2. $0 < q + r < 1$. Recall from Section 7.1.1 that $r < \pi^* < 1 - q$.

Case 2a. $\tilde{\pi} < r$. Then, $\tilde{V}(r) = \tilde{V}(r, L)$ and $\tilde{V}(1 - q) = \tilde{V}(1 - q, L)$. By $r < 1 - q$, we then have $\tilde{V}(1 - q) < \tilde{V}(r)$. Hence, in view of (13), $\tilde{\pi} > \frac{1}{2}$. As we have already shown in Theorem 2.2 that $\pi_1 \leq \frac{1}{2}$, we deduce $\tilde{\pi} \geq \pi_1$. On the other hand, $\tilde{\pi} < r < \pi^* = \pi_2$.

Case 2b. $r \leq \tilde{\pi} \leq 1 - q$. Then, $\tilde{V}(r) = \tilde{V}(r, R)$ and $\tilde{V}(1 - q) = \tilde{V}(1 - q, L)$. By (12) and (11),

$$\tilde{V}(r) = 1 + r\tilde{V}(1 - q) \quad \text{and} \quad \tilde{V}(1 - q) = 1 + q\tilde{V}(r),$$

yielding

$$\tilde{V}(r) = \frac{1 + r}{1 - qr} \quad \text{and} \quad \tilde{V}(1 - q) = \frac{1 + q}{1 - qr}.$$

Hence,

$$\tilde{\pi} = \frac{\tilde{V}(r)}{\tilde{V}(1 - q) + \tilde{V}(r)} = \frac{1 + r}{2 + q + r}.$$

Since $\pi_2 = \pi^* = \frac{r}{q+r}$, by using that $q \leq r$, a simple calculation shows that $\tilde{\pi} \leq \pi_2$.

Now we argue that $\tilde{\pi} \geq \pi_1$. If $\pi_1 \leq r$ then we are immediately ready as $\pi_1 \leq r \leq \tilde{\pi}$. So, assume that $\pi_1 > r$. Then, as we know from Section 7.1.1, Case (i),

$$\pi_1 = \frac{q(1 + r)}{(q + r)(1 + q)}.$$

By using that $q \leq r$, a simple calculation shows that $\tilde{\pi} \geq \pi_1$.

Case 2c. $1 - q < \tilde{\pi}$. We show that this case cannot occur. Assume that $q > 0$, otherwise we are done. Since $r < 1 - q$, the assumption $1 - q < \tilde{\pi}$ yields $\tilde{V}(r) = \tilde{V}(r, R)$ and $\tilde{V}(1 - q) = \tilde{V}(1 - q, R)$. By (12)

$$\tilde{V}(r) = 1 + r\tilde{V}(1 - q) \quad \text{and} \quad \tilde{V}(1 - q) = 1 + (1 - q)\tilde{V}(1 - q),$$

yielding

$$\tilde{V}(1 - q) = \frac{1}{q} \quad \text{and} \quad \tilde{V}(r) = 1 + \frac{r}{q}.$$

Hence,

$$\tilde{\pi} = \frac{\tilde{V}(r)}{\tilde{V}(1 - q) + \tilde{V}(r)} = \frac{q + r}{1 + q + r}.$$

Since $\tilde{\pi} > 1 - q$, we have

$$\frac{q + r}{1 + q + r} > 1 - q,$$

thus $q(q + r) > 1 - q$. By using $q + r < 1$, this implies $q > \frac{1}{2}$. As $r \geq q$, the inequality $q > \frac{1}{2}$ yields $q + r > 1$, which is a contradiction.

Case 3. $q + r = 1$. As $1 - q = r$, it follows from (13) that $\tilde{\pi} = \frac{1}{2}$. We know from Theorem 2.2 that $\pi_1 \leq \frac{1}{2}$ and $\pi_2 \geq \frac{1}{2}$, so we are done.

Case 4. $1 < q + r < 2$. Note that $q > 0$ and $r > 0$ must hold. Recall from Section 7.1.3 that $1 - q < \pi^* < r$.

Case 4a. $\tilde{\pi} < 1 - q$. Then, $\tilde{V}(1 - q) = \tilde{V}(1 - q, L)$ and $\tilde{V}(r) = \tilde{V}(r, L)$. By (11),

$$\tilde{V}(1 - q) = 1 + q\tilde{V}(r) \quad \text{and} \quad \tilde{V}(r) = 1 + (1 - r)\tilde{V}(r),$$

yielding

$$\tilde{V}(r) = \frac{1}{r} \quad \text{and} \quad \tilde{V}(1 - q) = 1 + \frac{q}{r}.$$

Hence,

$$\tilde{\pi} = \frac{\tilde{V}(r)}{\tilde{V}(1 - q) + \tilde{V}(r)} = \frac{1}{1 + q + r}. \tag{14}$$

We will now show that $\pi_1 < 1 - q$. As we know from Section 7.1.3, this is equivalent to showing that

$$(1 - q + q^2)r - 2q^2 + q^3 > 0.$$

As

$$\begin{aligned} (1 - q + q^2)r - 2q^2 + q^3 &= (r - qr - q^2) + q^2r - q^2 + q^3 \\ &= (r - qr - q^2) + q^2(q + r - 1) \end{aligned}$$

we must prove that

$$(r - qr - q^2) + q^2(q + r - 1) > 0. \tag{15}$$

The assumption $\tilde{\pi} < 1 - q$ together with (14) yields $(1 - q)(1 + q + r) > 1$, thus

$$r - qr - q^2 > 0. \quad (16)$$

By (16) and $q + r > 1$, we have shown (15). Hence, $\pi_1 < 1 - q$, and Case (ii) of Section 7.1.3 applies. Thus,

$$\pi_1 = \frac{1 - r}{(q + r)(1 - q) + 1 - r}. \quad (17)$$

Based on (14) and (17), with the assumption $q \leq r$, an easy calculation shows that $\tilde{\pi} \geq \pi_1$. On the other hand, $\tilde{\pi} < 1 - q < \pi^* = \pi_2$.

Case 4b. $1 - q \leq \tilde{\pi} \leq r$. Then, $\tilde{V}(1 - q) = \tilde{V}(1 - q, R)$ and $\tilde{V}(r) = \tilde{V}(r, L)$. By (12) and (11),

$$\tilde{V}(1 - q) = 1 + (1 - q)\tilde{V}(1 - q) \quad \text{and} \quad \tilde{V}(r) = 1 + (1 - r)\tilde{V}(r),$$

yielding

$$\tilde{V}(1 - q) = \frac{1}{q} \quad \text{and} \quad \tilde{V}(r) = \frac{1}{r}.$$

Hence,

$$\tilde{\pi} = \frac{\tilde{V}(r)}{\tilde{V}(1 - q) + \tilde{V}(r)} = \frac{q}{q + r}.$$

Since $\pi_2 = \pi^* = \frac{r}{q + r}$, by using that $q \leq r$, we may conclude that $\tilde{\pi} \leq \pi_2$.

Now we will argue that $\tilde{\pi} \geq \pi_1$. If $\pi_1 < 1 - q$ then we are immediately ready as $\pi_1 < 1 - q \leq \tilde{\pi}$. So, assume that $\pi_1 \geq 1 - q$. Then, as we know from Section 7.1.3, Case (i),

$$\pi_1 = \frac{q(1 - r)}{(q + r)(1 - q)}.$$

By using that $q \leq r$, we obtain

$$\pi_1 \leq \frac{q}{q + r} = \tilde{\pi}.$$

Case 4c. $r < \tilde{\pi}$. We show that this case cannot occur. The assumption $r < \tilde{\pi}$ yields $\tilde{V}(1 - q) = \tilde{V}(1 - q, R)$ and $\tilde{V}(r) = \tilde{V}(r, R)$. By (12)

$$\tilde{V}(1 - q) = 1 + (1 - q)\tilde{V}(1 - q) \quad \text{and} \quad \tilde{V}(r) = 1 + r\tilde{V}(1 - q),$$

yielding

$$\tilde{V}(1 - q) = \frac{1}{q} \quad \text{and} \quad \tilde{V}(r) = 1 + \frac{r}{q}.$$

Hence,

$$\tilde{\pi} = \frac{\tilde{V}(r)}{\tilde{V}(1 - q) + \tilde{V}(r)} = \frac{q + r}{1 + q + r}.$$

The assumption $\tilde{\pi} > r$ yields $q + r > r(1 + q + r)$, thus $q > r(q + r)$. As $q + r > 1$, this implies $q > r$, which is a contradiction.

Case 5. $q + r = 2$. Recall from Section 6.5 that $\pi_1 = \pi_2 = \frac{1}{2}$. Obviously, $q + r = 2$ yields $q = r = 1$. By using $\tilde{V}(1 - q) = \tilde{V}(0) = \tilde{V}(0, R) = 1$ and $\tilde{V}(r) = \tilde{V}(1) = \tilde{V}(1, L) = 1$, we obtain $\tilde{\pi} = \frac{1}{2}$.

8.3 Comparison of Values

We finally compare the value \tilde{V} in the model without waiting with the value V in the model with waiting. Clearly, $\tilde{V} \geq V$ always. We will now investigate the situations where the difference between the two is minimal and maximal, respectively.

It is intuitive that when $q = r = 0$ (Case 1) or $q = r = 1$ (Case 5), we have $\pi_1 = \pi_2 = \tilde{\pi} = \frac{1}{2}$ and hence $\tilde{V} = V$. In those cases, namely, there is no reason for the agent to wait. Hence, optimal strategies in the model with waiting and the optimal strategies in the model without waiting coincide.

Now, consider the situation where $q = 0, r = 1$ (a special case of Case 3). In this case, the agent would like to wait whenever $p \in (0, 1)$, so the waiting region is maximal here. Now suppose that $\bar{p} = \frac{1}{2}$. Then, $\tilde{V}(\frac{1}{2}) = \frac{3}{2}$, since by searching a location at $t = 1$ the agent will find the object immediately with probability $\frac{1}{2}$, and otherwise will find the object for sure at location L next period. On the other hand, in the model with waiting we have $V(\frac{1}{2}) = 1$, since by waiting at $t = 1$ the agent will find the object for sure at location L next period. We conjecture that the difference between the two values can never be more than $\frac{1}{2}$, so that this would be a case where the difference is maximal.

9 Multiple Agents

Now, let's assume that there are n agents searching for the same object. Similarly as in the original model, there are three possible actions in each period for every agent: searching L , searching R and waiting (0). At every period, an agent observes the actions chosen by his opponents in the previous period, and knows whether the object has been found or not. As usual, the object is initially at L with probability \bar{p} , and at R with probability $1 - \bar{p}$; and it moves according to a finite Markov process with transition probabilities q and r .

This extension makes the problem a more strategic one, since now the agents not only compete against time but also against each other in finding the object.⁷ If agent i is the first to find the object, possibly at the same time as some of his opponents⁸, then his utility would be equal to some prize P (for finding the object first) minus his total searching cost. If he does not find the object first, he will not receive the prize P but still incurs the cost of searching. Kan (1977), among others, considered two natural objectives in a search for a moving target problem: (i) maximizing the probability of finding the object in a given time, (ii) minimizing the expected cost for finding the object. These two objectives are, in a sense, embedded in the problem of each agent in our multiple agents extension. Since only the agents who find the object first receive the prize, each agent wants to maximize the probability of finding

⁷Nakai (1986) analyzed a different search game with two searchers and n locations, in which the object does not move. Two searchers with exponential detection functions compete against each other for a quicker detection of the object. Different prior beliefs over locations are allowed. In the case of identical prior beliefs, he showed that both players have the same equilibrium strategy even though their detection rates are different.

⁸ The situation changes if the prize is divided amongst the agents who find the object first.

the object in a given period. On the other hand, each agent wants to minimize his own expected searching cost.

Recall that in the one-agent case, a strategy was simply a list $\gamma = (\gamma_t)_{t \in \mathbb{N}}$ of actions. This definition is no longer suitable for the multiple agents case. The reason is that agent i is not certain about the actions of others, and therefore does not know in advance what action to take at period t . A strategy σ_i for agent i should therefore specify for every period t and *every possible sequence* h_{t-1} of past actions some action $\sigma_{i,t}(h_{t-1})$. Note that the probability p_t that the object is at L at period t is determined by h_{t-1} . A strategy σ_i is called *simple* if $\sigma_{i,t}(h_{t-1})$ only depends on t and p_t . In particular, if σ_i consists of a fixed sequence of actions (as in the one-agent case) then σ_i is simple.

Now, take a profile $(\sigma_1, \dots, \sigma_n)$ of strategies, one for each agent i . This profile is called a *Nash equilibrium* if no agent i can improve his expected utility by unilaterally changing his strategy, given that his opponents play according to $(\sigma_1, \dots, \sigma_n)$. The profile is called a *subgame perfect equilibrium* if at any period t and given any history h_{t-1} of past actions, the players' continuation strategies in $(\sigma_1, \dots, \sigma_n)$ constitute a Nash equilibrium.

We first show that, if $P \geq 2$, then typically all subgame perfect equilibria are symmetric, within the class of simple and strictly loss-free strategies.

Theorem 9.1 *Suppose that $P \geq 2$. Let p_t denote the probability that at period t the object is at L . Let $(\sigma_1, \dots, \sigma_n)$ be a subgame perfect equilibrium where σ_i for each agent i is simple and strictly loss-free. If $p_t \neq \frac{1}{2}$ at every period t after any past play, then $(\sigma_1, \dots, \sigma_n)$ must be symmetric.*

Proof. Suppose, contrary to what we want to prove, that $(\sigma_1, \dots, \sigma_n)$ is an asymmetric subgame perfect equilibrium consisting of simple and loss-free strategies. We analyze three different cases here.

(i) Suppose that at some period t , at least one agent searches L and some other agent, say agent i , searches R . Assume without loss of generality that $p_t > \frac{1}{2}$. Agent i 's expected utility from searching R at period t is

$$u_i^t = (1 - p_t)(P - 1) + p_t(-1),$$

whereas his expected utility from searching L and always waiting afterwards is

$$\tilde{u}_i^t = p_t(P - 1) + (1 - p_t)(-1) > (1 - p_t)(P - 1) + p_t(-1) = u_i^t$$

since $p_t > \frac{1}{2}$. Therefore, agent i would prefer searching L and waiting thereafter to searching R at period t , which is a contradiction.

(ii) Suppose that at some period t , at least one agent searches L , some other agent, say agent i , waits, and nobody searches R . Since the strategy of the agent that searches L is strictly loss-free, we have $p_t P - 1 > 0$. Agent i 's expected utility from waiting at period t is

$$u_i^t = p_t \cdot 0 + (1 - p_t)u_i^{t+1},$$

where u_i^{t+1} denotes the expected utility from period $t+1$ onwards. On the other hand, his expected utility from searching L at period t and copying his original strategy from period $t+1$ onwards would be

$$\tilde{u}_i^t = p_t(P-1) + (1-p_t)(u_i^{t+1} - 1) = p_t \cdot P - 1 + (1-p_t)u_i^{t+1} > u_i^t.$$

Here we used that the players all use simple strategies. Therefore, agent i would be better off by searching L at period t and copying his original strategy from period $t+1$ onwards, which is a contradiction.

(iii) Suppose that at some period t , at least one agent searches R , some other agent, say agent i , waits, and nobody searches L . This case is similar to case (ii). ■

Now, we show that, if $P \geq 2$, there is a symmetric subgame perfect equilibrium in which every agent implements his optimal one-person strategy, given an optimal strategy exists.

Theorem 9.2 *Suppose that $P \geq 2$, and $q = r$ or $q + r \geq 1$. Take a one-person optimal strategy σ from Theorem 2.2. Then, there is a symmetric subgame perfect equilibrium in which every agent uses σ .*

Proof. Suppose that the game has reached period t , and that the object has not yet been found. Let p_t be the probability that the object is at L . Let us focus on agent i , and let V_i^{t+1} be the highest utility that agent i could possibly obtain from period $t+1$ on, if his opponents play according to σ . We use the following two steps.

Step 1. Assume that σ tells agent i to search a location at period t , say location L . We show that it cannot be better to wait or to search the other location instead. We know from Theorem 2.2 that $p_t \geq \frac{1}{2}$. By searching L at period t , and acting optimally from period $t+1$ on, agent i 's expected utility from period t on would be

$$\begin{aligned} u_i &= p_t(P-1) + (1-p_t)(V_i^{t+1} - 1) \\ &= p_tP + (1-p_t)V_i^{t+1} - 1. \end{aligned}$$

If agent i would search R instead at period t , his expected utility from period t on would be

$$(1-p_t)(P-1) + p_t(-1) = (1-p_t)P - 1 \leq p_tP - 1 \leq u_i.$$

Finally, if agent i would wait at period t , then his expected utility from period t on would be at most

$$p_t0 + (1-p_t)V_i^{t+1} \leq u_i$$

since $p_tP - 1 \geq 0$. Hence, at period t it is a best reply for agent i to search L .

Step 2. We now show that for agent i it is optimal to use strategy σ if the others do so as well. Suppose that τ is a strategy for agent i which, at every period t , specifies an optimal action, given that the others play according to σ , and given that player i would obtain V_i^{t+1} if the object would not be found at period t . Similarly to Lemma 5.1, part (1), it follows that from every period t on agent i 's continuation

strategy in τ is a best reply against his opponents' continuation strategies in σ .⁹ By step 1, we may assume that, whenever σ prescribes to search a location, then τ prescribes to search the same location as well. So, the only difference could be that at some period σ prescribes to wait whereas τ prescribes to search a location. This implies that player i 's expected searching cost by using τ is exactly the same as it would be in the one-person case. Obviously, player i 's expected searching cost by using σ is also the same as in the one-person case. Since τ cannot yield a lower expected searching cost than σ in the one-person case, it can also not yield a lower expected searching cost than σ if the other agents play σ . If agent i plays σ , he will with probability 1 be the first to find the object. Therefore, τ cannot be a better reply for agent i than σ . So, it is optimal for agent i to play σ if the others play σ as well. Hence, if every agent searches according to σ this is a subgame perfect equilibrium. ■

In the above theorem, if only ε -optimal strategies exist, then they form a subgame perfect ε -equilibrium in a similar way.

We now show that, if $P \geq 2$, there exists a subgame perfect equilibrium in which every agent implements the one-person myopic strategy (i.e., search the location with the higher current probability of containing the object).

Theorem 9.3 *Suppose that $P \geq 2$. Take a strategy σ that always searches the location with the highest current probability of containing the object (if both locations are equally probable, then choose either one). Then, there is a symmetric subgame perfect equilibrium in which every agent uses σ .*

Proof. Suppose that the game has reached period t , and that the object has not yet been found. Let p_t be the probability that the object is at L , and let $w_t = \max\{p_t, 1 - p_t\} \geq \frac{1}{2}$. Let us focus on agent i , and let V_i^{t+1} be the highest utility that agent i could possibly obtain from period $t + 1$ on, if his opponents play according to σ . Then, by searching the location with the highest current probability at period t , and acting optimally from period $t + 1$ on, agent i 's expected utility from period t on would be

$$\begin{aligned} u_i &= w_t(P - 1) + (1 - w_t)(V_i^{t+1} - 1) \\ &= w_tP + (1 - w_t)V_i^{t+1} - 1. \end{aligned}$$

If agent i would search the other location instead at period t , his expected utility from period t on would be

$$(1 - w_t)(P - 1) + w_t(-1) = (1 - w_t)P - 1 \leq w_tP - 1 \leq u_i.$$

Finally, if agent i would wait at period t , then his expected utility from period t on would be at most

$$w_t0 + (1 - w_t)V_i^{t+1} \leq u_i$$

⁹ The strategies of the opponents of player i will guarantee that the object will eventually be found with probability 1. Hence the condition with α in Lemma 3 plays no role here.

since $w_t P - 1 \geq 0$. Hence, at every period t it is a best reply for agent i to search the location with the highest current probability of containing the object if the opponents do so as well. Similarly to Lemma 5.1, part (1), it then follows that from every period t on agent i 's continuation strategy in σ is a best reply against his opponents' continuation strategies. So, it is a subgame perfect equilibrium. ■

For two strategy profiles $(\sigma_1, \dots, \sigma_n)$ and (τ_1, \dots, τ_n) , we say that the first *Pareto dominates* the second if for every agent i his expected utility in the first is at least as high as in the second. It is easily seen that the subgame perfect equilibrium from Theorem 9.2 Pareto dominates the one from Theorem 9.3. In fact, the following corollary proves a more general result.

Theorem 9.4 *Suppose that $P \geq 2$. Then, the subgame perfect equilibrium $(\sigma_1, \dots, \sigma_n)$ in which every agent uses his one-person optimal strategy Pareto dominates all other symmetric subgame perfect equilibria.*

Proof. Take another symmetric subgame perfect equilibrium (τ_1, \dots, τ_n) . The expected utility that any agent i would get under (τ_1, \dots, τ_n) is equal to his expected utility from playing τ_i in the one-person case. Similarly, the expected utility that any agent i would get under $(\sigma_1, \dots, \sigma_n)$ is equal to his expected payoff from playing σ_i in the one-person case. Since σ_i is optimal in the one-person case, agent i 's expected utility under $(\sigma_1, \dots, \sigma_n)$ will be at least as high as his expected utility under (τ_1, \dots, τ_n) . ■

The results above are no longer true if $P < 2$. Consider, for instance, the situation where the initial probability \bar{p} is $\frac{1}{2}$, and the transition probabilities q and r are $\frac{1}{2}$ as well. If all players would follow one of the strategies above, then the expected searching cost for every agent would be exactly 2, and hence it would be better for every agent to wait forever instead.

If the prize P is chosen large enough then we can prove the following result: Take a strategy σ which finds the object with probability 1 in every subgame, and never prescribes to wait. Then, the symmetric strategy profile in which every agent plays σ will be a subgame perfect equilibrium if P is chosen large enough. Namely, if σ prescribes to search a location, then searching the other location or waiting will not be optimal if P is large, since it would yield the risk of not finding the object first. In particular, it will be a subgame perfect equilibrium if everybody always searches the location with the lowest probability of containing the object. Note that this behavior was never optimal in the one-agent case.

10 Discussion

1. Assumption on the Division of P in the Case of Simultaneous Discovery

Note that in Section 9 on multiple agents, we assumed that if multiple agents find the object at the same time, each receives the prize, P . This assumption

keeps the case of multiple agents similar to the case of a single agent. If we assume an equal division of the prize in the case of simultaneous discovery, there exist values of initial probabilities for which there is no symmetric Nash equilibrium. This is due to the fact that in this case agents have a higher incentive to search another location than the others. Therefore, the object will be found with probability 1 in period 1 in those equilibria. Hence, dividing the prize equally in the case of simultaneous discovery leads to a more cost-efficient outcome for the planner compared to the case in which the prize is not divided. However, if there are many agents then it might be necessary to increase the lower bound on P to make agents search at all.

2. *Inefficiency of the Equilibria in the Case of Multiple Agents*

From the planner's point of view, the most efficient strategy profile would be one in which at $t = 1$ both locations are searched, which would lead to the immediate discovery of the object. This, however, cannot be achieved by a subgame perfect equilibrium in general, except for some cases when the initial probability \bar{p} equals $\frac{1}{2}$.¹⁰

3. *Possibility of Overlooking*

Weber (1986) and MacPhee & Jordan (1995) analyzed the search for a moving target problem with a single agent and without waiting in continuous time and discrete time, respectively. Their models take the possibility of overlooking into account. Thus, in their models, even if the object is at location L (or R) and the agent searches that location, he may overlook the object with some positive probability. This possibility affects the updating process. Now, when the agent searches a location and cannot find the object, he should take into account the possibility that the object was in fact at the location he searched. In the multiple agents case it can happen, for instance, that even though both locations are searched by different agents, the game will not end with probability 1.

4. *Existence of Optimal Strategies*

In the model without the option to wait, optimal strategies always exist (cf. Pollock, 1970, and MacPhee & Jordan, 1995). However, as we mentioned in the analysis above, when waiting is a (costless) option, an optimal strategy may not exist in the non-oscillating case (i.e., $0 < q + r < 1$). Nevertheless, we proved that for every $\varepsilon > 0$, there exists an ε -optimal strategy.

5. *More than Two Locations*

The case of more than two locations is much more difficult than the case of two locations. Even in the case of three locations, one should provide two-dimensional areas of probability distributions that describe an optimal search

¹⁰ Consider the completely symmetric situation where $\bar{p} = \frac{1}{2}$ and $q = r$. Suppose there are 4 agents. Then, if agents 1 and 2 search location L while agents 3 and 4 search location R , this will yield a subgame perfect equilibrium in which the object is found immediately.

strategy instead of thresholds on a line. (See Nakai (1973) and MacPhee and Jordan (1995) for some partial results for the model without waiting.)

Appendix: 3 Sub-cases with respect to the values of π^* and $\bar{\pi}$

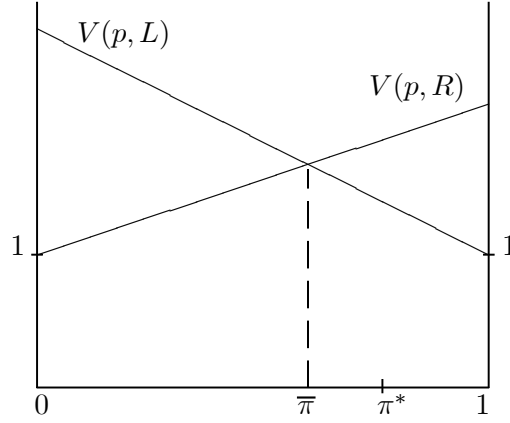


Figure 9: The functions $V(p, L)$ and $V(p, R)$ when $\bar{\pi} < \pi^*$ (i.e. $q < r$)

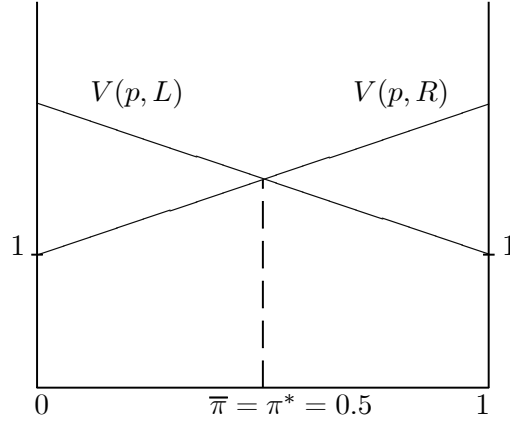


Figure 10: The functions $V(p, L)$ and $V(p, R)$ when $\bar{\pi} = \pi^*$ (i.e. $q = r$)

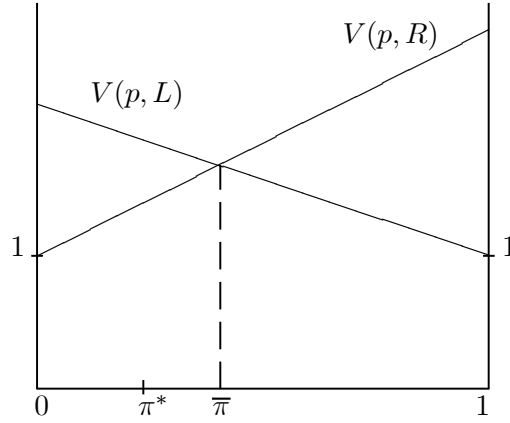


Figure 11: The functions $V(p, L)$ and $V(p, R)$ when $\bar{\pi} > \pi^*$ (i.e. $q > r$)

References

- [1] Assaf, D. & Sharlin-Bilitzky, A. (1994) "Dynamic Search for a Moving Target", *Journal of Applied Probability*, 31(2), pp. 438 – 457.
- [2] Kan, Y.C. (1977) "Optimal Search of a Moving Target", *Operations Research*, 25(5), pp. 864 – 870.
- [3] MacPhee, I.M. & Jordan, B.P. (1995) "Optimal Search for a Moving Target", *Probability in the Engineering and Informational Sciences* **9**, pp. 159 – 182.
- [4] Nakai, T. (1973) "A Model of Search for a Target Among Three Boxes: Some Special Cases", *Journal of Operations Research Society of Japan*, 16, pp. 151 - 162.
- [5] Nakai, T. (1986) "A Search Game with One Object and Two Searchers", *Journal of Applied Probability*, 23(3), pp. 696 - 707.
- [6] Pollock, S.M. (1970) "A Simple Model of Search for a Moving Target", *Operations Research*, 18(5), pp. 883 – 903.
- [7] Pursiheimo, U. (1978) "On the Optimal Search for a Moving Target in Discrete Space", In *Lecture Notes in Control and Information Sciences 6: Optimization Techniques*, pp. 177 – 185, Springer, Berlin-Heidelberg-New York.
- [8] Ross, S (1983) *Introduction to Stochastic Dynamic Programming*, Academic Press, New York.
- [9] Weber, R.R. (1986) "Optimal Search for a Randomly Moving Object", *Journal of Applied Probability*, 23(3), pp. 708 – 717.